

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>6</sup> :		A1	(11) International Publication Number: WO 99/03990
C12N 15/11, 15/63, 15/00, 15/12, A61K 38/17, C07K 16/00, C12P 21/02, C12Q 1/68, G01N 33/68			(43) International Publication Date: 28 January 1999 (28.01.99)
(21) International Application Number: PCT/US98/14613		(71) Applicant (for all designated States except US): HUMAN GENOME SCIENCES, INC. [US/US]; 9410 Key West Avenue, Rockville, MD 20850 (US).	
(22) International Filing Date: 15 July 1998 (15.07.98)		(72) Inventors; and	
(30) Priority Data:		(73) Inventors/Applicants (for US only): RUBEN, Steven, M. [US/US]; 18528 Heritage Hills Drive, Olney, MD 20832 (US). ROSEN, Craig, A. [US/US]; 22400 Rolling Hill Road, Laytonsville, MD 20882 (US). YOUNG, Paul, E. [US/US]; 122 Beckwith Street, Gaithersburg, MD 20878 (US). GREENE, John, M. [US/US]; 872 Diamond Drive, Gaithersburg, MD 20878 (US). NI, Jian [CN/US]; 5502 Manorfield Road, Rockville, MD 20853 (US). FENG, Ping [CN/US]; 4 Relda Court, Gaithersburg, MD 20878 (US). FLORENCE, Kimberly, A. [US/US]; 12805 Atlantic Avenue, Rockville, MD 20851 (US). HU, Jing-Shan [CN/US]; 1247 Lakeside Drive #3034, Sunnyvale, CA 94086 (US). FERRIE, Ann, M. [US/US]; 13203 L. Astoria Hill Court, Germantown, MD 20874 (US). YU, Guo-Liang [CN/US]; 1714C Marina Court, San Mateo, CA 94403 (US). DUAN, Roxanne [US/US]; 5515 Northfield Road, Bethesda, MD 20817 (US). FOUAD, Janet [SY/US]; #202, 140 High Street, Westerly, RI 02891 (US).	
60/052,661	16 July 1997 (16.07.97)	US	(74) Agents: BROOKES, A., Anders et al.; Human Genome Sciences, Inc., 9410 Key West Avenue, Rockville, MD 20850 (US).
60/052,872	16 July 1997 (16.07.97)	US	(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).
60/052,871	16 July 1997 (16.07.97)	US	
60/052,874	16 July 1997 (16.07.97)	US	
60/052,873	16 July 1997 (16.07.97)	US	
60/052,870	16 July 1997 (16.07.97)	US	
60/052,875	16 July 1997 (16.07.97)	US	
60/053,440	22 July 1997 (22.07.97)	US	
60/053,441	22 July 1997 (22.07.97)	US	
60/053,442	22 July 1997 (22.07.97)	US	
60/055,726	18 August 1997 (18.08.97)	US	
60/055,724	18 August 1997 (18.08.97)	US	
60/056,359	18 August 1997 (18.08.97)	US	
60/055,725	18 August 1997 (18.08.97)	US	
60/055,985	18 August 1997 (18.08.97)	US	
60/055,952	18 August 1997 (18.08.97)	US	
60/055,946	18 August 1997 (18.08.97)	US	
60/055,683	18 August 1997 (18.08.97)	US	
60/055,989	18 August 1997 (18.08.97)	US	
60/056,361	18 August 1997 (18.08.97)	US	
Published <i>With international search report.</i>			

(54) Title: 64 HUMAN SECRETED PROTEINS

(57) Abstract

The present invention relates to 64 novel human secreted proteins and isolated nucleic acids containing the coding regions of the genes encoding such proteins. Also provided are vectors, host cells, antibodies, and recombinant methods for producing human secreted proteins. The invention further relates to diagnostic and therapeutic methods useful for diagnosing and treating disorders related to these novel human secreted proteins.

***FOR THE PURPOSES OF INFORMATION ONLY***

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

## 64 Human Secreted Proteins

### *Field of the Invention*

This invention relates to newly identified polynucleotides and the polypeptides encoded by these polynucleotides, uses of such polynucleotides and polypeptides, and  
5 their production.

### *Background of the Invention*

Unlike bacterium, which exist as a single compartment surrounded by a membrane, human cells and other eucaryotes are subdivided by membranes into many functionally distinct compartments. Each membrane-bounded compartment, or  
10 organelle, contains different proteins essential for the function of the organelle. The cell uses "sorting signals," which are amino acid motifs located within the protein, to target proteins to particular cellular organelles.

One type of sorting signal, called a signal sequence, a signal peptide, or a leader sequence, directs a class of proteins to an organelle called the endoplasmic reticulum (ER). The ER separates the membrane-bounded proteins from all other types of  
15 proteins. Once localized to the ER, both groups of proteins can be further directed to another organelle called the Golgi apparatus. Here, the Golgi distributes the proteins to vesicles, including secretory vesicles, the cell membrane, lysosomes, and the other organelles.

20 Proteins targeted to the ER by a signal sequence can be released into the extracellular space as a secreted protein. For example, vesicles containing secreted proteins can fuse with the cell membrane and release their contents into the extracellular space - a process called exocytosis. Exocytosis can occur constitutively or after receipt of a triggering signal. In the latter case, the proteins are stored in secretory vesicles (or  
25 secretory granules) until exocytosis is triggered. Similarly, proteins residing on the cell membrane can also be secreted into the extracellular space by proteolytic cleavage of a "linker" holding the protein to the membrane.

Despite the great progress made in recent years, only a small number of genes encoding human secreted proteins have been identified. These secreted proteins include  
30 the commercially valuable human insulin, interferon, Factor VIII, human growth hormone, tissue plasminogen activator, and erythropoietin. Thus, in light of the pervasive role of secreted proteins in human physiology, a need exists for identifying and characterizing novel human secreted proteins and the genes that encode them. This knowledge will allow one to detect, to treat, and to prevent medical disorders by using  
35 secreted proteins or the genes that encode them.

### *Summary of the Invention*

The present invention relates to novel polynucleotides and the encoded polypeptides. Moreover, the present invention relates to vectors, host cells, antibodies, 5 and recombinant methods for producing the polypeptides and polynucleotides. Also provided are diagnostic methods for detecting disorders related to the polypeptides, and therapeutic methods for treating such disorders. The invention further relates to screening methods for identifying binding partners of the polypeptides.

10

### *Detailed Description*

#### Definitions

The following definitions are provided to facilitate understanding of certain terms used throughout this specification.

In the present invention, "isolated" refers to material removed from its original 15 environment (e.g., the natural environment if it is naturally occurring), and thus is altered "by the hand of man" from its natural state. For example, an isolated polynucleotide could be part of a vector or a composition of matter, or could be contained within a cell, and still be "isolated" because that vector, composition of matter, or particular cell is not the original environment of the polynucleotide.

20 In the present invention, a "secreted" protein refers to those proteins capable of being directed to the ER, secretory vesicles, or the extracellular space as a result of a signal sequence, as well as those proteins released into the extracellular space without necessarily containing a signal sequence. If the secreted protein is released into the extracellular space, the secreted protein can undergo extracellular processing to produce 25 a "mature" protein. Release into the extracellular space can occur by many mechanisms, including exocytosis and proteolytic cleavage.

As used herein, a "polynucleotide" refers to a molecule having a nucleic acid 30 sequence contained in SEQ ID NO:X or the cDNA contained within the clone deposited with the ATCC. For example, the polynucleotide can contain the nucleotide sequence of the full length cDNA sequence, including the 5' and 3' untranslated sequences, the coding region, with or without the signal sequence, the secreted protein coding region, as well as fragments, epitopes, domains, and variants of the nucleic acid sequence. Moreover, as used herein, a "polypeptide" refers to a molecule having the translated 35 amino acid sequence generated from the polynucleotide as broadly defined.

In the present invention, the full length sequence identified as SEQ ID NO:X was often generated by overlapping sequences contained in multiple clones (contig

analysis). A representative clone containing all or most of the sequence for SEQ ID NO:X was deposited with the American Type Culture Collection ("ATCC"). As shown in Table 1, each clone is identified by a cDNA Clone ID (Identifier) and the ATCC Deposit Number. The ATCC is located at 10801 University Boulevard, 5 Manassas, Virginia 20110-2209, USA. The ATCC deposit was made pursuant to the terms of the Budapest Treaty on the international recognition of the deposit of microorganisms for purposes of patent procedure.

A "polynucleotide" of the present invention also includes those polynucleotides capable of hybridizing, under stringent hybridization conditions, to sequences contained 10 in SEQ ID NO:X, the complement thereof, or the cDNA within the clone deposited with the ATCC. "Stringent hybridization conditions" refers to an overnight incubation at 42° C in a solution comprising 50% formamide, 5x SSC (750 mM NaCl, 75 mM sodium citrate), 50 mM sodium phosphate (pH 7.6), 5x Denhardt's solution, 10% dextran sulfate, and 20 µg/ml denatured, sheared salmon sperm DNA, followed by washing the 15 filters in 0.1x SSC at about 65°C.

Also contemplated are nucleic acid molecules that hybridize to the polynucleotides of the present invention at lower stringency hybridization conditions. Changes in the stringency of hybridization and signal detection are primarily accomplished through the manipulation of formamide concentration (lower percentages 20 of formamide result in lowered stringency); salt conditions, or temperature. For example, lower stringency conditions include an overnight incubation at 37°C in a solution comprising 6X SSPE (20X SSPE = 3M NaCl; 0.2M NaH<sub>2</sub>PO<sub>4</sub>; 0.02M EDTA, pH 7.4), 0.5% SDS, 30% formamide, 100 µg/ml salmon sperm blocking DNA; followed by washes at 50°C with 1XSSPE, 0.1% SDS. In addition, to achieve even 25 lower stringency, washes performed following stringent hybridization can be done at higher salt concentrations (e.g. 5X SSC).

Note that variations in the above conditions may be accomplished through the inclusion and/or substitution of alternate blocking reagents used to suppress background in hybridization experiments. Typical blocking reagents include 30 Denhardt's reagent, BLOTO, heparin, denatured salmon sperm DNA, and commercially available proprietary formulations. The inclusion of specific blocking reagents may require modification of the hybridization conditions described above, due to problems with compatibility.

Of course, a polynucleotide which hybridizes only to polyA+ sequences (such 35 as any 3' terminal polyA+ tract of a cDNA shown in the sequence listing), or to a

complementary stretch of T (or U) residues, would not be included in the definition of "polynucleotide," since such a polynucleotide would hybridize to any nucleic acid molecule containing a poly (A) stretch or the complement thereof (e.g., practically any double-stranded cDNA clone).

The polynucleotide of the present invention can be composed of any polyribonucleotide or polydeoxyribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. For example, polynucleotides can be composed of single- and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. In addition, the polynucleotide can be composed of triple-stranded regions comprising RNA or DNA or both RNA and DNA. A polynucleotide may also contain one or more modified bases or DNA or RNA backbones modified for stability or for other reasons. "Modified" bases include, for example, tritylated bases and unusual bases such as inosine. A variety of modifications can be made to DNA and RNA; thus, "polynucleotide" embraces chemically, enzymatically, or metabolically modified forms.

The polypeptide of the present invention can be composed of amino acids joined to each other by peptide bonds or modified peptide bonds, i.e., peptide isosteres, and may contain amino acids other than the 20 gene-encoded amino acids. The polypeptides may be modified by either natural processes, such as posttranslational processing, or by chemical modification techniques which are well known in the art. Such modifications are well described in basic texts and in more detailed monographs, as well as in a voluminous research literature. Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. It will be appreciated that the same type of modification may be present in the same or varying degrees at several sites in a given polypeptide. Also, a given polypeptide may contain many types of modifications. Polypeptides may be branched, for example, as a result of ubiquitination, and they may be cyclic, with or without branching. Cyclic, branched, and branched cyclic polypeptides may result from posttranslation natural processes or may be made by synthetic methods. Modifications include acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphatidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cysteine,

formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, pegylation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination. (See, for instance, PROTEINS - STRUCTURE AND MOLECULAR PROPERTIES, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York (1993); POSTTRANSLATIONAL COVALENT MODIFICATION OF PROTEINS, B. C. Johnson, Ed., Academic Press, New York, pgs. 1-12 (1983); Seifter et al., Meth Enzymol 182:626-646 (1990); Rattan et al., Ann NY Acad Sci 663:48-62 (1992).)

"SEQ ID NO:X" refers to a polynucleotide sequence while "SEQ ID NO:Y" refers to a polypeptide sequence, both sequences identified by an integer specified in Table 1.

"A polypeptide having biological activity" refers to polypeptides exhibiting activity similar, but not necessarily identical to, an activity of a polypeptide of the present invention, including mature forms, as measured in a particular biological assay, with or without dose dependency. In the case where dose dependency does exist, it need not be identical to that of the polypeptide, but rather substantially similar to the dose-dependence in a given activity as compared to the polypeptide of the present invention (i.e., the candidate polypeptide will exhibit greater activity or not more than about 25-fold less and, preferably, not more than about tenfold less activity, and most preferably, not more than about three-fold less activity relative to the polypeptide of the present invention.)

## 25 Polynucleotides and Polypeptides of the Invention

### FEATURES OF PROTEIN ENCODED BY GENE NO: 1

Preferred polypeptides encoded by this gene comprise the following amino acid sequence:

30 IRHELGCSWRFRAVKAAQGLFLSAPGPAARRCHGVVRCFSTCRALTA  
RCTGRVPWEACLYSSEPPLTETVARSVSWTCELALTCYAPRALSGAPVLCRHD  
V (SEQ ID NO:155). Also provided are polynucleotides encoding such polypeptides.

This gene is expressed in human substantia nigra tissue. Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are

not limited to, neurological disorders or abnormalities. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the neurological systems, expression of this 5 gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., brain and of the tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue 10 or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:83 as residues: Pro-30 to Leu-35.

The tissue distribution indicates that the protein products of this gene are useful for diagnosis and treatment of neurological disorders and abnormalities. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible 15 through sequence databases. Some of these sequences are related to SEQ ID NO:11 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides 20 comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 544 of SEQ ID NO:11, b is an integer of 15 to 558, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:11, and where the b is greater than or equal to a + 14.

## 25 FEATURES OF PROTEIN ENCODED BY GENE NO: 2

This gene is expressed primarily in breast and testes.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are 30 not limited to, tumors, particularly those of the breast or testes. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the breast and testes, expression of this gene at significantly higher or lower levels may be routinely detected in certain 35 tissues or cell types (e.g., reproductive tissue, and cancerous and wounded tissues) or bodily fluids (e.g., breast milk, seminal fluid, serum, plasma, urine, synovial fluid and

spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:84 as residues:

5 Ser-32 to His-37.

The tissue distribution indicates that the protein products of this gene are useful for diagnosis and treatment of disorders or abnormalities of breast and testes such as tumors of those tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences 10 are related to SEQ ID NO:12 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general 15 formula of a-b, where a is any integer between 1 to 701 of SEQ ID NO:12, b is an integer of 15 to 715, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:12, and where the b is greater than or equal to a + 14.

### FEATURES OF PROTEIN ENCODED BY GENE NO: 3

20 This gene is expressed in apoptotic T cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorder or abnormalities of T cells. Similarly, polypeptides and 25 antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., cells and tissues of the immune system, developmental tissue, and 30 cancerous and wounded tissues) or bodily fluids (e.g., lymph seminal fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in 35 SEQ ID NO:85 as residues: Met-1 to Glu-6, Leu-39 to Lys-46.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of disorders in T cells and other immune system disorders such as inflammation. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:13 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 824 of SEQ ID NO:13, b is an integer of 15 to 838, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:13, and where the b is greater than or equal to a + 14.

#### 15 FEATURES OF PROTEIN ENCODED BY GENE NO: 4

The translation product of this gene shares sequence homology with ubiquitin-conjugating enzyme (UCE) involved in selective protein degradation. Based on the sequence similarity, the translation product of this gene is expected to share biological activities with UCE proteins. Such activities are known in the art and described elsewhere herein. Preferred polypeptides encoded by this gene comprise the following amino acid sequence:

FLAIHFPTDFPLKPPKVAFTRMYFPNSNSNGSTCLDILWSQWSPAL (SEQ ID NO: 156). Also provided are polynucleotides encoding such polypeptides.

This gene is expressed primarily in testes.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders in testes, particularly cell cycle disorders, (e.g. testes tumor). Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the testes and reproductive systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., reproductive tissues, and cancerous and wounded tissues) or bodily fluids (e.g., seminal fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene

expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to ubiquitin-conjugating enzyme indicates that the protein product of this gene is useful for diagnosis and treatment of disorders in testes and reproductive system such as tumors, as well as the treatment of tumors of other origins. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:14 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 499 of SEQ ID NO:14, b is an integer of 15 to 513, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:14, and where the b is greater than or equal to a + 14.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 5

This gene is expressed primarily in testes.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders and abnormalities in the testes. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., reproductive tissue, and cancerous and wounded tissues) or bodily fluids (e.g., seminal fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, (i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder). Preferred epitopes include those comprising a sequence shown in SEQ ID NO:87 as residues: Ser-22 to Thr-32, Pro-37 to Ser-42.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of disorders in testes and the reproductive system. Many polynucleotide sequences, such as EST sequences,

are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:15 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 698 of SEQ ID NO:15, b is an integer of 15 to 712, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:15, and where the b is greater than or equal to a + 14.

10

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 6

This gene is expressed primarily in thymus, activated monocytes and spleen. Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders and abnormalities of immune function and hematopoiesis, e.g. leukemia. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and hematopoietic systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., endocrine tissue cell and tissue of the immune system, and haematopoietic tissue, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that the protein products of this gene are useful for diagnosis and treatment of immunological and hematopoietic disorders such as leukemia. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:16 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 638 of SEQ ID NO:16, b is an integer of 15 to

652, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:16, and where the b is greater than or equal to a + 14.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 7

5 This gene is expressed in T cells.  
Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, T cell dysfunction. Similarly, polypeptides and antibodies directed to  
10 these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., cells and tissue of the immune system, and cancerous and wounded  
15 tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

20 The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of disorders in T cells and immune systems. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:17 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically  
25 excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 728 of SEQ ID NO:17, b is an integer of 15 to 742, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:17, and where the b is greater than or equal to a + 14.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 8

This gene is expressed primarily in pleural cancer and to a less extent in T cells.  
Therefore, polynucleotides and polypeptides of the invention are useful as  
35 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are

not limited to, pleural cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the pleural system, expression of this gene at 5 significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., cells and tissue of the immune system, pleural tissue and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in 10 healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:90 as residues: Ser-30 to Tyr-37.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of pleural cancer.

15 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:18 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.

20 Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1205 of SEQ ID NO:18, b is an integer of 15 to 1219, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:18, and where the b is greater than or equal to a + 14.

25

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 9

This gene is expressed in endothelial cells that shares the same origin as hematopoietic cells and in spleen and liver which are hematopoietic tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as 30 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders of endothelial cells or hematopoiesis. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of 35 disorders of the above tissues or cells, particularly of the endothelial and hematopoietic systems, expression of this gene at significantly higher or lower levels may be routinely

detected in certain tissues or cell types (e.g., endothelial tissue, haematopoietic cells and tissue, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of disorders in endothelial or hematopoietic systems. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:19 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 860 of SEQ ID NO:19, b is an integer of 15 to 874, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:19, and where the b is greater than or equal to a + 14.

## 20 FEATURES OF PROTEIN ENCODED BY GENE NO: 10

This gene is expressed primarily in breast lymph node and to a lesser extent in melanocytes.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, metastatic melanoma. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the lymphatic system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., reproductive tissue, cells and tissue of the immune system, cancerous and wounded tissues) or bodily fluids (e.g., lymph, breast milk, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution of this gene in melanocytes and lymph node indicates that the protein product of this gene is useful for the diagnosis and treatment of metastatic melanoma involving lymphatic tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of 5 these sequences are related to SEQ ID NO:20 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the 10 general formula of a-b, where a is any integer between 1 to 450 of SEQ ID NO:20, b is an integer of 15 to 464, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:20, and where the b is greater than or equal to a + 14.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 11

15 This gene is believed to reside on chromosome 2. Therefore, this gene is useful in linkage analysis as a marker for chromosome 2.

This gene is expressed primarily in infant brain and to a lesser extent in fetal liver/spleen.

Therefore, polynucleotides and polypeptides of the invention are useful as 20 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurological disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above 25 tissues or cells, particularly of the central nervous system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., neural tissue, cells and tissue of the immune system, developing tissue, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken 30 from an individual having such a disorder, relative to the standard gene expression level, (i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder). Preferred epitopes include those comprising a sequence shown in SEQ ID NO:93 as residues: Tyr-59 to Gln-68, His-84 to Leu-90, Ser-105 to Asn-110, Leu-112 to Pro-118.

35 The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of neurological

disorders. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:21 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 623 of SEQ ID NO:21, b is an integer of 15 to 637, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:21, and where the b is greater than or equal to a + 14.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 12

This gene is expressed primarily in adipose tissue.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, metabolic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the endocrine system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., adipose tissue, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for regulation of fat metabolism and treatment of obesity. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:22 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 738 of SEQ ID NO:22, b is an integer of 15 to

752, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:22, and where the b is greater than or equal to a + 14.

### FEATURES OF PROTEIN ENCODED BY GENE NO: 13

5       The translation product of this gene shares sequence homology with NADH:ubiquinone oxidoreductase, the first enzyme in the respiratory electron transport chain of mitochondria.

This gene is expressed primarily in HSC172 cells and to a lesser extent in pineal gland.

10      Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, jet lag. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential 15 identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the endocrine system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., endocrine tissue, metabolic tissue, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or 20 another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level; i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:95 as residues: Thr-30 to Val-38, Glu-49 to Ile-54.

25      The tissue distribution and homology to NADH:ubiquinone oxidoreductase indicates that polynucleotides and polypeptides corresponding to this gene are useful for minimizing the negative effects of travel across time zones by altering the body's circadian clock. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are 30 related to SEQ ID NO:23 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of 35 a-b, where a is any integer between 1 to 478 of SEQ ID NO:23, b is an integer of 15 to

492, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:23, and where the b is greater than or equal to a + 14.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 14

5 This gene is expressed primarily in synovial IL-1/TNF stimulated cells. Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, arthritis. Similarly, polypeptides and antibodies directed to these 10 polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the skeletal system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., skeletal tissue and cancerous and wounded tissues) or bodily fluids (e.g., 15 lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides 20 corresponding to this gene are useful for treatment of degenerative conditions of joints, including arthritis. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:24 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically 25 excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 518 of SEQ ID NO:24, b is an integer of 15 to 532, where both a and b correspond to the positions of nucleotide 30 residues shown in SEQ ID NO:24, and where the b is greater than or equal to a + 14.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 15

This gene is expressed primarily in 12 Week Old Early Stage Human. Therefore, polynucleotides and polypeptides of the invention are useful as 35 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are

not limited to, developmental disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the fetal systems, expression of this gene at significantly 5 higher or lower levels may be routinely detected in certain tissues or cell types (e.g., developing tissue, and cancerous and wounded tissues) or bodily fluids (e.g., amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an 10 individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of some developmental disorders. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences 15 are related to SEQ ID NO:25 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general 20 formula of  $a-b$ , where  $a$  is any integer between 1 to 906 of SEQ ID NO:25,  $b$  is an integer of 15 to 920, where both  $a$  and  $b$  correspond to the positions of nucleotide residues shown in SEQ ID NO:25, and where the  $b$  is greater than or equal to  $a + 14$ .

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 16

25 This gene is expressed primarily in thymus and to a lesser extent in neutrophils. Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune disorders. Similarly, polypeptides and antibodies directed to 30 these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., cells and tissue of the immune system, and cancerous and wounded 35 tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder,

relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of some immune disorders including lupus and other disorders involving thymic dysfunction. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:26 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 903 of SEQ ID NO:26, b is an integer of 15 to 917, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:26, and where the b is greater than or equal to a + 14.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 17

This gene is expressed primarily in fibrosarcoma and to a lesser extent in IL1 and IPS induced neutrophils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, fibrosarcoma. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., connective tissue, and cells and tissue of the immune system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:99 as residues: Gly-6 to Pro-11.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of fibrosarcoma or

other immune disorders. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:27 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 648 of SEQ ID NO:27, b is an integer of 15 to 662, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:27, and where the b is greater than or equal to a + 14.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 18

Preferred polypeptides comprise the following amino acid sequence:

MLLTPHFNVANPQNLLAGLWLENEHSFTLMAPERARTHHCQPEERKVLFCLFP  
15 IVPNSQAQVQPPQMPPFCCAAAKEKTQEQLQEPLGSQCPDTCPNSLC (SEQ ID NO: 157). Polynucleotides encoding such polypeptides are also provided.

This gene is expressed primarily in jurkat T-Cells in S phase, and to a lesser extent in IL-1 and LPS induced neutrophils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., cells and tissue of the immune system, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:100 as residues: Lys-97 to Gln-106, Gln-112 to Pro-118, Pro-123 to Lys-130, Arg-153 to Gly-158.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of immune disorders related to jurkat T-cells and induced neutrophils. Many polynucleotide sequences, such

as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:28 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To 5 list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 685 of SEQ ID NO:28, b is an integer of 15 to 699, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:28, and where the b is greater 10 than or equal to a + 14.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 19

The translation product of this gene shares sequence homology with FK506-binding protein FKBP-12 and FKBP13, members of a family of proteins which bind 15 the immunosuppressant drugs FK506 and rapamycin. The homology to FK506 binding proteins indicates that the protein encoded by this gene has similar activity to the known FK506 BP family members. Such activity may be assayed according to methods known in the art and described elsewhere herein. Preferred polypeptides encoded by this gene comprise the following amino acid sequence:

20 MRLFLWNAVLTLFVTSLIGALIPEPEVK  
IEVLQKPFICHRKTKGGDMLVHYEGYLEKD GSLFHSTHKHNNGQPIWFTLGI  
LEALKGWDQGLKGMCVGEKRKLIPPALGYGKEGKGKIPPESTLIFNIDLLEIR  
NGPRSHESFQEMDLNDDWKLSKDEVKAYLKKEFEKHGAVVNESHHDALVEDI  
FDKEDEDKDGFISAREFTYKHDEL (SEQ ID NO: 158) or comprise a mature form  
25 of the foregoing polypeptide having the following amino acid sequence:  
EVKIEVLQKPFICHRKTKGGDMLVHYEGYLEKD GSLFHSTHKHNNGQPIWF  
TLGILEALKGWDQGLKGMCVGEKRKLIPPALGYGKEGKGKIPPESTLIFNDL  
LEIRNGPRSHESFQEMDLNDDWKLSKDEVKAYLKKEFEKHGAVVNESHHDAL  
VEDIFDKEDEDKDGFISAREFTYKHDEL (SEQ ID NO: 159). Polynucleotides  
30 encoding these polypeptides are also provided. Particularly preferred is the following  
polynucleotide sequence:

ATGAGGCTTTCTTGTGGAACGCGGTCTTGACTCTGTCGTCACITCT  
TTGATTGGGGCTTGATCCCTGAACCAAGAAGTAAAATTGAAGTTCTCCAGA  
AGCCATTCATCTGCCATCGCAAGACCAAAGGAGGGATTGATGTTGGTCC  
35 ACTATGAAGGCTACTTAGAAAAGGACGGCTCCTTATTCACTCCACTCACAA  
ACATAACAATGGTCAGCCCATTGGTTACCTGGGCATCCTGGAGGCTCTC

AAAGGTTGGGACCAGGGCTTGAAAGGAATGTGTAGGAGAGAAGAGAAA  
GCTCATTCCTCCTGCTCTGGCTATGGAAAAGAAGGAAAAGGTAAAATT  
CCCCCAGAAAGTACACTGATATTAAATATTGATCTCCTGGAGATTGAAATG  
GACCAAGATCCCATGAATCATTCCAAGAAATGGATCTTAATGATGACTGGAA  
5 ACTCTCTAAAGATGAGGTTAAAGCATATTAAAGAAGGAGTTGAAAAACAT  
GGTCGGTGGTGAATGAAAGTCATCATGATGCTTGGTGGAGGATATTTT  
GATAAAGAAGATGAAGACAAAGATGGTTATATCTGCCAGAGAATTACAT  
ATAAACACGATGAG TTA (SEQ ID NO:160), and the portion of it, nucleotide  
residues 76 to the 3' end, which encode the mature form shown above.

10 This gene is expressed primarily in fetal heart, ovary, and thymus.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune dysfunction such as autoimmune disorders. Similarly,

15 polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., cells and tissue of the immune system, and  
20 cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

25 The tissue distribution and homology to FK506-binding protein FKBP-12 and FKBP13 indicates that polynucleotides and polypeptides corresponding to this gene are useful for identifying immunosuppressant drugs, and may be used in combination with immunosuppressant drugs for therapeutic purposes in the treatment of autoimmune diseases and organ/tissue transplant rejection. Many polynucleotide sequences, such as  
30 EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:29 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the  
35 present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1623 of

SEQ ID NO:29, b is an integer of 15 to 1637, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:29, and where the b is greater than or equal to a + 14.

## 5 FEATURES OF PROTEIN ENCODED BY GENE NO: 20

The protein product of this gene is believed to be the human homologue of yeast GAA1P which anchors proteins to cell surface membranes by glycosylphosphatidylinositols (GPIs). See, for example, Yeast Gaalp is required for attachment of a completed GPI anchor onto proteins (J Cell Biol. 1995 May; 129(3): 10 629-639. PMID: 7730400; UI: 95247814).

This gene is expressed primarily in primary breast cancer and to a lesser extent in amniotic cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, breast cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the metabolic system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., reproductive tissue, amniotic cells and tissue and cancerous and wounded tissues) or bodily fluids (e.g., breast milk, lymph, amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:102 as residues: Pro-113 to Met-123.

The tissue distribution and similarity to GAA1P indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of primary breast cancers. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:30 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general

formula of a-b, where a is any integer between 1 to 2128 of SEQ ID NO:30, b is an integer of 15 to 2142, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:30, and where the b is greater than or equal to a + 14.

## 5 FEATURES OF PROTEIN ENCODED BY GENE NO: 21

This gene is expressed primarily in activated T-cells and to a lesser extent in prostate cancer and HSC172 cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune dysfunction and/or prostate cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and metabolic systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., cells and tissue of the immune system, reproductive tissue, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, seminal fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:103 as residues: Ala-17 to Pro-26, Phe-92 to Tyr-97, Gly-104 to Glu-111.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of immune disorders and/or prostate cancer. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:31 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1550 of SEQ ID NO:31, b is an integer of 15 to 1564, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:31, and where the b is greater than or equal to a + 14.

**FEATURES OF PROTEIN ENCODED BY GENE NO: 22**

This gene maps to chromosome 12, and therefore, may be used as a marker in linkage analysis for chromosome 12.

5 This gene is expressed primarily in breast tissues and to a lesser extent in fetal spleen.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, breast cancer, lactation deficiencies, and other breast related diseases, 10 including breast cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the haemolymphoid and reproductive systems, expression of this gene at significantly higher or lower levels may be routinely detected 15 in certain tissues or cell types (e.g., breast tissue, cells and tissue of the immune system, and haemolymphoid tissue, cancerous and wounded tissues) or bodily fluids (e.g., lymph, breast milk, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level; i.e., the expression level in healthy tissue or 20 bodily fluid from an individual not having the disorder.

The tissue distribution in tumors of breast origins indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and intervention of these tumors, in addition to other tumors where expression has been indicated. Protein, as well as, antibodies directed against the protein may show utility as a tissue-specific 25 marker and/or immunotherapy target for the above listed tissues. Alternatively, The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are . useful for the treatment and diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. The uses include 30 bone marrow cell ex vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia. The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors 35 of various blood lineages, and in the differentiation and/or proliferation of various cell types. Many polynucleotide sequences, such as EST sequences, are publicly available

and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:32 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.

5 Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1617 of SEQ ID NO:32, b is an integer of 15 to 1631, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:32, and where the b is greater than or equal to a + 14.

10

### FEATURES OF PROTEIN ENCODED BY GENE NO: 23

The translation product of this gene shares sequence homology with mutant or polymorphic BRCA1 gene which is thought to be important in the diagnosis and therapy of human breast and ovarian cancer as a predisposing gene (See Genebank Accession No T17455). This gene maps to chromosome 18, and therefore, may be used 15 as a marker in linkage analysis for chromosome 18.

This gene is expressed primarily in primary dendritic cells. Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a 20 biological sample and for diagnosis of diseases and conditions which include, but are not limited to, breast, ovarian, and other reproductive related disorders and cancers. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell 25 type(s). For a number of disorders of the above tissues or cells, particularly of the haemolymphoid and reproductive systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., reproductive tissue, pleural tissue, breast tissue, and cancerous and wounded tissues) or bodily fluids (e.g., breast milk, lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a 30 disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:105 as residues: Pro-13 to Lys-18, Ala-50 to Leu-58.

The homology to the mutant or polymorphic BRCA1 gene indicates that 35 polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and intervention of breast and/or ovarian tumors, in addition to other tumors where

expression has been indicated. Protein, as well as, antibodies directed against the protein may show utility as a tissue-specific marker and/or immunotherapy target for the above listed tissues. Alternatively, the tissue distribution within dendritic cells indicates that the protein product of this gene is useful for the detection/treatment of neurodegenerative disease states and behavioral disorders such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses , autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:33 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 964 of SEQ ID NO:33, b is an integer of 15 to 978, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:33, and where the b is greater than or equal to a + 14.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 24

The translation product of this gene shares sequence homology with smaller hepatocellular oncoprotein which is thought to be important in protein synthesis (See Genebank Accession No. R07057).One embodiment of this gene comprises polypeptides of the following amino acid sequence: LRSVVQDHGQHGETPSLLKIQ @ (SEQ ID NO:161). An additional embodiment is the polynucleotides encoding these polypeptides.

This gene is expressed primarily in embryonic tissues and to a lesser extent in uterine cancer.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, hepatocellular tumors, and other disorders characterized by proliferating and/or developing tissues. Similarly, polypeptides and antibodies directed to these

polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the haemolymphoid system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., reproductive tissue, cells and tissue of the immune system, and cancerous and wounded tissues) or bodily fluids (e.g., amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:106 as residues: Asn-15 to Ser-20, Ile-32 to Asn-37.

The tissue distribution combine with the homology to a conserved human hepatocellular oncoprotein indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of cancer and other proliferative disorders. Expression within embryonic tissue and other cellular sources marked by proliferating cells indicates that this protein may play a role in the regulation of cellular division. Additionally, the expression also indicates that this protein may play a role in the proliferation, differentiation, and/or survival of hematopoietic cell lineages. In such an event, this gene may be useful in the treatment of lymphoproliferative disorders, and in the maintenance and differentiation of various hematopoietic lineages from early hematopoietic stem and committed progenitor cells. Similarly, embryonic development also involves decisions involving cell differentiation and/or apoptosis in pattern formation. Thus this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy.

Alternatively, the homology to a hepatocellular protein indicates that the protein product of this gene is useful for the detection and treatment of liver disorders and cancers (e.g. hepatoblastoma, jaundice, hepatitis, liver metabolic diseases and conditions that are attributable to the differentiation of hepatocyte progenitor cells). In addition the expression in fetus would suggest a useful role for the protein product in developmental abnormalities, fetal deficiencies, pre-natal disorders and various wound-healing models and/or tissue trauma. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:34 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or

more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 884 of SEQ ID NO:34, b is an integer of 15 to 898, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:34, and where the b is greater than or equal to a + 14.

5

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 25

The translation product of this gene shares sequence homology with a gene encoding for anti-heparanase activity which is thought to be important in inhibition of heparin or heparan sulphate degradation. Moreover, this gene was shown to have

10 homology to the human 3-oxo-5-beta-steroid 4-dehydrogenase, which is known to be important in metabolism since it catalyzes the reduction of delta(4) double bonds of bile acid intermediates and steroid hormones carrying the delta(4)-3-one structure in the A/B Cis configuration (See Genebank Accession No. Z28339) One embodiment of this gene comprises polypeptides of the following amino acid sequence:

15 MFYNFVRQLDTVSIIEHAGSKLKMTVGTKLTSGXGPRKSSQSGRIAASITDCQ  
QCKA @ (SEQ ID NO:162), and/or  
MEAAILPLWLLFLGPXPEVSFVPTVIFNLDFPACSLTVSSCLTKL @ (SEQ ID  
NO:163). An additional embodiment is the polynucleotides encoding these  
polypeptides.

20 This gene is expressed primarily in fetal tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, arthritis, metabolic disorders, as well as reproductive disorders.

25 Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the haemolymphoid system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g. reproductive, metabolic, 30 cancerous and wounded tissues) or bodily fluids (e.g., bile acid, amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

35 The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of cancer and other

proliferative disorders. Expression within embryonic tissue and other cellular sources marked by proliferating cells indicates that this protein may play a role in the regulation of cellular division. Additionally, the expression indicates that this protein may play a role in the proliferation, differentiation, and/or survival of hematopoietic cell lineages.

5 In such an event, this gene may be useful in the treatment of lymphoproliferative disorders, and in the maintenance and differentiation of various hematopoietic lineages from early hematopoietic stem and committed progenitor cells. Similarly, embryonic development also involves decisions involving cell differentiation and/or apoptosis in pattern formation. Thus this protein may also be involved in apoptosis or tissue

10 differentiation and could again be useful in cancer therapy. Alternatively, based upon its homology to a conserved human anti-heparanase gene, mutations of which are known to be important in the predisposition of arthritis, may suggest that this protein may also be important in the diagnosis or treatment of various autoimmune disorders such as rheumatoid arthritis, lupus, scleroderma, and dermatomyositis as well as dwarfism,

15 spinal deformation, and specific joint abnormalities as well as chondrodysplasias i.e., spondyloepiphyseal dysplasia congenita, familial osteoarthritis, Atelosteogenesis type II, metaphyseal chondrodysplasia type Schmid. Moreover, the homology to a conserved human metabolic gene may suggest that the protein product of this gene is useful for the diagnosis, prevention, and/or treatment of various metabolic disorders

20 such as Tay-Sachs disease, phenylketonuria, galactosemia, porphyrias, and Hurler's syndrome. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:35 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded

25 from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 740 of SEQ ID NO:35, b is an integer of 15 to 754, where both a and b correspond to the positions of nucleotide residues shown in

30 SEQ ID NO:35, and where the b is greater than or equal to a + 14.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 26

The translation product of this gene shares sequence homology with pulmonary hydrophobic surfactant-associated proteins which is thought to be important in useful for normalising pulmonary surface tension (See Genebank Accession No. N80643).

This gene is expressed primarily in embryonic tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, hyaline membrane disease, developmental and pulmonary disorders.

5     Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the respiratory system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., pulmonary tissue, developing 10 tissue and cancerous and wounded tissues) or bodily fluids (e.g., surfactant, amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

15     The tissue distribution and homology to pulmonary hydrophobic surfactant-associated proteins indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of such lung related diseases such as hyaline membrane disease which is often characteristic of premature infants - leading to significant pulmonary disorders throughout childhood. Alternatively, The tissue 20 distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of cancer and other proliferative disorders. Expression within embryonic tissue and other cellular sources marked by proliferating cells indicates that this protein may play a role in the regulation of cellular division. Additionally, the expression indicates that this protein may play a role in the 25 proliferation, differentiation, and/or survival of hematopoietic cell lineages. In such an event, this gene may be useful in the treatment of lymphoproliferative disorders, and in the maintenance and differentiation of various hematopoietic lineages from early hematopoietic stem and committed progenitor cells. Similarly, embryonic development also involves decisions involving cell differentiation and/or apoptosis in pattern 30 formation. Thus this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:36 and may have been publicly 35 available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the

present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 685 of SEQ ID NO:36, b is an integer of 15 to 699, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:36, and where the b is greater than or equal to a + 14.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 27

This gene is expressed primarily in fetal tissue and to a lesser extent in thymus. Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immunological disorders, particularly immunodeficiency. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the haemolymphoid system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., cells and tissue of the immune system, developing tissue, and cancerous and wounded tissues) or bodily fluids (e.g., amniotic fluid, lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:109 as residues: Pro-18 to Trp-24.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Expression of this gene product in thymus indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, and leukemia. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tumors and tissues. In addition, this gene product may have

commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:37 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 957 of SEQ ID NO:37, b is an integer of 15 to 971, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:37, and where the b is greater than or equal to a + 14.

15

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 28**

This gene is expressed primarily in small intestine and to a lesser extent in ulcerative colitis.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, ulcers, acid reflux and other gastrointestinal disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the digestive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., gastrointestinal tissue, and cancerous and wounded tissues) or bodily fluids (e.g., bile, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:110 as residues: Ile-2 to Ser-8, Gln-23 to Ser-31, Lys-61 to Lys-66, Lys-74 to Thr-79, Val-138 to Glu-160, Glu-178 to Thr-183.

35

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of digestive and

gastrointestinal problems such as acid reflux or ulcers. Alternatively, the tissue distribution may suggest that the protein product of this gene is useful for the diagnosis, prevention, and/or treatment of various metabolic disorders such as Tay-Sachs disease, phenylketonuria, galactosemia, porphyrias, and Hurler's syndrome.

- 5 Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:38 and may have been publicly available prior to conception of the present invention. Preferably,
- 10 such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 858 of SEQ ID NO:38, b is an integer of 15 to 872, where both a and b
- 15 correspond to the positions of nucleotide residues shown in SEQ ID NO:38, and where the b is greater than or equal to a + 14.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 29

The translation product of this gene was shown to have homology to catalase (See Genebank Accession No. pirlA55092|A55092) which is known to play an integral role in the oxidative prophylaxis in eukaryotic and mammalian cells. One embodiment of this gene comprises polypeptides of the following amino acid sequence:  
NHGHSCFLCEIVRSQFHTTYEPEA @ (SEQ ID NO:164), and/or  
SGRHRVELQLLFPLVRVNFEGLGVNHGSCFLCEIVRSQFHTTYEPEA @ (SEQ  
25 ID NO:165). An additional embodiment is the polynucleotides encoding these polypeptides. This gene maps to chromosome 3, and therefore, may be used as a marker in linkage analysis for chromosome 3.

This gene is expressed primarily in skin and to a lesser extent in human uterine cancer.

- 30 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, uterine cancer; fibrosis; melanoma. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes
- 35 for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the epidermis and/or reproductive system,

expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., integumentary tissue, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, 5 relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution combined with the homology to catalase indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and/or diagnosis of human uterine cancer and potentially, all cancers in general 10 (particularly of the epidermis) due to the fact that diminished catalase activity has been shown to lead to significant cellular damage which could predispose cancer. Likewise, this gene and/or its protein product may be useful in the treatment and/or diagnosis of other disorders of the skin, such as fibrosis or in wound healing. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or 15 immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:39 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To 20 list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 594 of SEQ ID NO:39, b is an integer of 15 to 608, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:39, and where the b is greater 25 than or equal to a + 14.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 30

This gene is expressed primarily in fetal liver/spleen and bone marrow stromal cells, and to a lesser extent in neural tissues (brain; spinal cord) and cancers 30 (glioblastoma; chondrosarcoma).

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cancers, particularly of glial cells and cartilage; hematopoietic and other 35 immune disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the

tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hematopoietic/immune system and CNS, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., hematopoietic cells and tissue, and cancerous and wounded tissues) or 5 bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:112 as residues: Arg-2 to Asp-7.

10 The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the expansion, proliferation, and/or differentiation of hematopoietic cells. This gene is expressed predominantly at sites of hematopoiesis, such as fetal liver and bone marrow, and thus may control the proliferation and/or differentiation of hematopoietic stem and progenitor cells. Thus it  
15 could be used for chemoprotection, or for the production of specific blood cell lineages, as well as the amplification of stem cells. Additionally, its expression in neural cells of the brain and spinal cord suggest that it may also play a role in the maintenance and differentiation of neuronal stem cells, or in the treatment of neurological disorders, such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome,  
20 schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses , autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked  
25 disorders, or disorders of the cardiovascular system. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Finally, its observed expression in various cancers suggest that it may play a role in the control of cell proliferation. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through  
30 sequence databases. Some of these sequences are related to SEQ ID NO:40 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a  
35 nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 841 of SEQ ID NO:40, b is an integer of 15 to 855, where both a and b

correspond to the positions of nucleotide residues shown in SEQ ID NO:40, and where the b is greater than or equal to a + 14.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 31

5 This gene is expressed primarily in placenta, bone marrow, and fetal liver/spleen and to a lesser extent in brain and CNS.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are 10 not limited to, hematopoietic disorders; placental insufficiency; neurological disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system and/or CNS, expression of this gene at significantly higher or lower 15 levels may be routinely detected in certain tissues or cell types (e.g., hematopoietic cells and tissue, cells and tissue of the immune system, and neurological tissue, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression 20 level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:113 as residues: Arg-53 to Gln-58.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the proliferation, maintenance, and/or 25 differentiation of hematopoietic cells. In addition, it may be involved in the maintenance and establishment of the vasculature, and may play a role in the regulation of angiogenesis. Thus, it may play a role in the establishment and/or maintenance of tumors. Expression in the CNS indicates that this gene product may also be useful in the treatment of neurodegenerative disorders, such as Alzheimers Disease, Parkinsons 30 Disease, Huntingtons Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses , autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the 35 developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. Protein, as well as, antibodies directed against the protein may show utility as a

tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:41 and may have been publicly available prior to conception of the present invention.

5 Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1028 of SEQ ID NO:41, b is an integer of 15 to 1042, where  
10 both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:41; and where the b is greater than or equal to a + 14.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 32

The translation product of this gene was shown to have homology to the ras-related protein RABB from Dictyostelium discoideum (See Genebank Accession No.P34142). Members of the Ras family of proteins are known to be essential to normal cell cycle control and mutations of which have been shown to lead to the predisposition of cancer.

This gene is expressed primarily in Jurkat T cells.  
20 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, T cell lymphomas; defects in immune surveillance and/or T cell activation. Similarly, polypeptides and antibodies directed to these polypeptides are  
25 useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., cells and tissue of the immune system, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken  
30 from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution combined with its homology to a Ras-related protein indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment of defects in immune surveillance and cancer. Expression of this gene

product in T cells indicates that it represents a secreted protein that may regulate T cells in an autocrine fashion, thereby impacting on their ability to recognize antigen and become activated, or may be involved in immune modulation. Likewise, this may represent a secreted protein made by T cells that affects other hematopoietic cells and regulates their proliferation and/or differentiation. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:42 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 688 of SEQ ID NO:42, b is an integer of 15 to 702, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:42, and where the b is greater than or equal to a + 14.

### FEATURES OF PROTEIN ENCODED BY GENE NO: 33

The translation product of this gene shares sequence homology with a known human spliceosome associated protein (See Genebank Accession No. AA523942) as well as a yeast protein (ORF YBR173c [Saccharomyces cerevisiae]). Preferred polypeptides comprise the following amino acid sequence:  
MNARGLGSELKDSIPVTELSASGPFES  
HDLLRKGFSCVKNELLPSHPLELSEKNFQLNQDKMNFSTLRNIQGLFAPLKLQ  
MEFKAVQQVQRLPFLSSNLSLDVLRGNDTIGFEDILNDPSQSEVMGEPHLMV  
EYKLGLL (SEQ ID NO:166). Also preferred are the polynucleotides encoding these polypeptides.

This gene is expressed primarily in bone marrow and stromal cells, and to a lesser extent in T cells and peripheral blood cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, hematopoietic disorders; leukemias; defects in immune surveillance; defects in T cell activation; inflammation; bacterial infections, and other disorders characterized by immunodeficiency. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential

identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., cells and tissue of the immune system, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:115 as residues: Lys-16 to Thr-24.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the proliferation, differentiation, and/or activation of hematopoietic cell lineages. Such treatments could involve chemoprotection or expansion of either progenitor cells or specific mature blood lineages. Likewise, this gene product could be involved in immune modulation, or in affecting T cell activation and antigen recognition. Alternatively, the protein product of this gene is useful for the treatment and diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. The uses include bone marrow cell ex vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia. The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:43 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 628 of SEQ ID NO:43, b is an integer of 15 to 642, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:43, and where the b is greater than or equal to a + 14.

## FEATURES OF PROTEIN ENCODED BY GENE NO: 34

This gene maps to chromosome 4, and therefore, may be used as a marker in linkage analysis for chromosome 4.

5 This gene is expressed primarily in placenta and brain, and to a lesser extent in specific tumors and cancers (kidney, colorectal, colon, osteoclastoma).

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are  
10 not limited to, tumors of the kidney, colon, and bone; neurological disorders; defects of the vasculature. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the vasculature and CNS, expression of this gene at significantly higher  
15 or lower levels may be routinely detected in certain tissues or cell types (e.g., reproductive tissue, cells and tissue of the immune system, neural tissue, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e.,  
20 the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment of neurological disorders, such as Alzheimers or schizophrenia. Expression of this gene in the placenta may indicate a role in fetal development, or may simply be a hallmark of expression in the vasculature.  
25 Expression of this gene product in endothelial cells may indicate secretion of the protein product into the circulation, where it may have effects on circulating blood cells, or on tissues at distant locations. At such sites, it may control cellular proliferation and/or differentiation. In addition, expression in a variety of tumors indicates that this gene  
30 product may play a role in cellular proliferation. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:44 and may have been publicly available prior to  
35 conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related

sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1205 of SEQ ID NO:44, b is an integer of 15 to 1219, where both a and b correspond to the positions of 5 nucleotide residues shown in SEQ ID NO:44, and where the b is greater than or equal to a + 14.

### FEATURES OF PROTEIN ENCODED BY GENE NO: 35

This gene is expressed primarily in macrophages, treated with GM-CSF.

10 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune deficiencies; susceptibility to bacterial infections; improper stimulation of lymphocyte pools. Similarly, polypeptides and antibodies directed to 15 these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., cells and tissue of the immune system, and cancerous and wounded 20 tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:117 as residues: Arg-48 to Asn-56.

25 The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the regulation and/or stimulation of hematopoietic cells, particularly the circulating lymphocytes. Macrophages are one of the front lines of immune defense, particularly against bacterial pathogens, and are able to secrete cytokines and proteins that affect other blood cells. Thus, this gene product 30 may be important in regulation of the immune system, in activation of hematopoietic cells, such as T cells; and may play a role in antigen recognition. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence 35 databases. Some of these sequences are related to SEQ ID NO:45 and may have been publicly available prior to conception of the present invention. Preferably, such related

polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 423 of SEQ ID NO:45, b is an integer of 15 to 437, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:45, and where the b is greater than or equal to a + 14.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 36

This gene is expressed primarily in endothelial cells. Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, inflammation and circulatory system disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the circulatory system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., endothelial cells, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of inflammatory disorders involving endothelial cells, such as sepsis, inflammatory bowel diseases, psoriasis, and rheumatoid arthritis as well as atherosclerosis, which can lead to strokes and heart attacks. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:46 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of

a-b, where a is any integer between 1 to 519 of SEQ ID NO:46, b is an integer of 15 to 533; where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:46, and where the b is greater than or equal to a + 14.

## 5 FEATURES OF PROTEIN ENCODED BY GENE NO: 37

The translation product of this gene was shown to have homology to the human transducin(beta)-like 1. protein (See Genebank Accession No P38262).

This gene is expressed primarily in fetal lung and to a lesser extent in adult lung and breast.

10 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, respiratory system diseases, and/or disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological 15 probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the respiratory system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., pulmonary tissue, and cancerous and wounded tissues) or bodily fluids (e.g., pulmonary surfactant, serum, plasma, urine, synovial fluid and 20 spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

25 The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and disease affecting the lung such as lung cancer, emphysema, pulmonary edema, asthma, cystic fibrosis, and pulmonary emboli. Alternatively, the protein product of this gene may be useful in the detection, treatment, and/or prevention of various inflammatory conditions afflicting endothelial tissue such as the vasculature and cardiovascular systems. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or 30 immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:47 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To 35 list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence

described by the general formula of a-b, where a is any integer between 1 to 1835 of SEQ ID NO:47, b is an integer of 15 to 1849, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:47, and where the b is greater than or equal to a + 14.

5

### FEATURES OF PROTEIN ENCODED BY GENE NO: 38

This gene is expressed primarily in prostate cancer and to a lesser extent in osteoblasts.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, prostate cancer, osteoporosis. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., skeletal tissue, reproductive tissue, and cancerous and wounded tissues) or bodily fluids (e.g., seminal fluid, lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in tumors of prostate origins indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and intervention of these tumors, in addition to other tissues where expression has been indicated. Protein, as well as, antibodies directed against the protein may show utility as a tissue-specific marker and/or immunotherapy target for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:48 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 912 of SEQ ID NO:48, b is an integer of 15 to 926, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:48, and where the b is greater than or equal to a + 14.

## FEATURES OF PROTEIN ENCODED BY GENE NO: 39

The translation product of this gene was shown to have homology to the mdkk-1 protein from *Mus musculus* (See Genebank Accession No. gil2736292 (AF030433)) which has been shown to be important as a possible homeobox gene inducer specific to head development. One embodiment of this gene comprises polypeptides of the following amino acid sequence: GGNKYQTIDNYQPYP @ (SEQ ID NO:167), PLLGVSATLNSVLNSNAIKN @ (SEQ ID NO:168), and/or GSAVSAAPGILYPG. An additional embodiment is the polynucleotides encoding these polypeptides (SEQ ID NO:169).

This gene is expressed primarily in placenta and to a lesser extent in smooth muscle.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, developmental disorders of the fetus, and/or reproductive disorders, particularly of the female. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of developing tissues in the fetus, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., reproductive tissue, and cancerous and wounded tissues) or bodily fluids (e.g., amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution combined with the homology to a suspected homeobox domain inducer indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of developmental disorders, particularly those involving hematopoiesis and pattern formation in the embryo. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:49 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To

list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1579 of SEQ ID NO:49, b is an integer of 15 to 1593, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:49, and where the b is greater than or equal to a + 14.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 40

When tested against sensory neuron cell lines (PC12), supernatants removed from cells containing this gene activated the early growth response gene 1 pathway (EGR1). Thus, it is likely that this gene activates neuronal cells through the Jaks-STAT signal transduction pathway. EGR1 (early growth response gene 1) is a separate signal transduction pathway from Jaks-STAT, genes containing the EGR1 promoter are induced in various tissues and cell types upon activation, leading the cells to undergo differentiation and proliferation.

This gene is expressed primarily in neutrophils and to a lesser extent in pancreatic carcinoma.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, inflammation, gastrointestinal, and immune disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., cells and tissue of the immune system, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:122 as residues: Pro-8 to Ala-16.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Expression of this gene product in tonsils indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of

potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin,  
5 the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, and leukemia. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tumors and tissues. In addition, this gene product may have  
10 commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through  
15 sequence databases. Some of these sequences are related to SEQ ID NO:50 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a  
20 nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 964 of SEQ ID NO:50, b is an integer of 15 to 978, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:50, and where the b is greater than or equal to a + 14.

## 25 FEATURES OF PROTEIN ENCODED BY GENE NO: 41

This gene is expressed primarily in breast lymph node and to a lesser extent in anergic T-cell.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a  
30 biological sample and for diagnosis of diseases and conditions which include, but are not limited to, inflammation, and immunodeficiency disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of  
35 this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., cells and tissue of the immune system, and cancerous and

wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

5       The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Expression of this gene product in T-cells indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product  
10      may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune  
15      deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tumors and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in  
20      the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:51 and may have been  
25      publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 419 of  
30      SEQ ID NO:51, b is an integer of 15 to 433, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:51, and where the b is greater than or equal to a + 14.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 42

35      This gene is expressed primarily in synovium.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, inflammation and diseases of the joints. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system and connective tissue, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., skeletal tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of inflammatory disorders, particularly those involving the joints and skeletal system, such as rheumatoid arthritis and in particular the connective tissues (e.g. trauma, tendonitis, chondromalacia). Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:52 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 847 of SEQ ID NO:52, b is an integer of 15 to 861, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:52, and where the b is greater than or equal to a + 14.

30

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 43

This gene maps to chromosome 5, and therefore, may be used as a marker in linkage analysis for chromosome 5.

This gene is expressed primarily in synovium.  
Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a

biological sample and for diagnosis of diseases and conditions which include, but are not limited to, diseases of the joints and connective tissues. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders 5 of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., skeletal tissue, and cancerous and wounded tissues) or bodily fluids (e.g. serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene 10 expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of inflammatory joint disorders as well as suggest a role in the detection and treatment of disorders and 15 conditions affecting the skeletal system, in particular the connective tissues (e.g. arthritis, trauma, tendonitis, chondromalacia and inflammation). Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence 20 databases. Some of these sequences are related to SEQ ID NO:53 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence 25 described by the general formula of a-b, where a is any integer between 1 to 496 of SEQ ID NO:53, b is an integer of 15 to 510, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:53, and where the b is greater than or equal to a + 14.

### 30 FEATURES OF PROTEIN ENCODED BY GENE NO: 44

This gene is expressed primarily in activated T cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are 35 not limited to, inflammation and other immune system disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing

immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., cells and tissue of the immune system, and  
5 cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID  
10 NO:126 as residues: Met-1 to Lys-7.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Expression of this gene product in T-cells indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of  
15 potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also  
20 used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tumors and tissues. In addition, this gene product may have commercial utility in  
25 the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence  
30 databases. Some of these sequences are related to SEQ ID NO:54 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence  
35 described by the general formula of a-b, where a is any integer between 1 to 295 of SEQ ID NO:54, b is an integer of 15 to 309, where both a and b correspond to the

positions of nucleotide residues shown in SEQ ID NO:54, and where the b is greater than or equal to a + 14.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 45

5 This gene is expressed primarily in placenta, liver, lung, endometrial stromal cell and embryo.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions such as cancers, 10 immunodeficiency and autoimmune diseases, as well as reproductive and hepatic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may 15 be routinely detected in certain tissues or cell types (e.g., reproductive tissue, and cancerous and wounded tissues) or bodily fluids (e.g., amniotic fluid, bile, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not 20 having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection and treatment of liver disorders and cancers (e.g. hepatoblastoma, jaundice, hepatitis, liver metabolic diseases and conditions that are attributable to the differentiation of hepatocyte progenitor cells). In 25 addition the expression in fetus would suggest a useful role for the protein product in developmental abnormalities, fetal deficiencies, pre-natal disorders and various wound-healing models and/or tissue trauma. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly 30 available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:55 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more 35 polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1571 of SEQ ID NO:55, b is an integer of 15

to 1585, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:55, and where the b is greater than or equal to a + 14.

### FEATURES OF PROTEIN ENCODED BY GENE NO: 46

5 This gene is expressed primarily in placenta, amniotic cells and adrenal gland tumor.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions such as cancers, as well 10 as various reproductive and endocrine disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell 15 types (e.g., endocrine tissue, reproductive tissue, and cancerous and wounded tissues) or bodily fluids (e.g., amniotic fluid, serum, plasma; urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include 20 those comprising a sequence shown in SEQ ID NO:128 as residues: Cys-52 to Val-61.

The tissue distribution within placenta and amniotic tissues indicates that polynucleotides and polypeptides corresponding to this gene useful for the treatment and diagnosis of reproductive disorders, particularly those of the female reproductive system, including cancer and tissues characteristic of the developing embryo. In 25 addition, the tissue distribution within the adrenal gland indicates that the protein product of this gene is useful for the detection, treatment, and/or prevention of various endocrine disorders and cancers, particularly Addison's disease, Cushing's Syndrome, and disorders and/or cancers of the pancreas (e.g. diabetes mellitus), adrenal cortex, ovaries, pituitary (e.g., hyper-, hypopituitarism), thyroid (e.g. hyper-, 30 hypothyroidism), parathyroid (e.g. hyper-, hypoparathyroidism), hypothalamus, and testes. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:56 and 35 may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the

present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 860 of SEQ ID NO:56, b is an integer of 15 to 874, where 5 both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:56, and where the b is greater than or equal to a + 14.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 47

This gene is expressed primarily in melanocyte, melanoma, dendritic cells and 10 fetal brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, skin disorders, particularly melanoma, as well as neurodegenerative 15 disorders and cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell 20 types (e.g., neural tissue, integumentary tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those 25 comprising a sequence shown in SEQ ID NO:129 as residues: Lys-76 to Gly-81.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment, diagnosis, and/or prevention of various skin disorders including congenital disorders (i.e. nevi, moles, freckles, Mongolian spots, hemangiomas, port-wine syndrome), integumentary tumors (i.e. 30 keratoses, Bowen's disease, basal cell carcinoma, squamous cell carcinoma, malignant melanoma, Paget's disease, mycosis fungoides, and Kaposi's sarcoma), injuries and inflammation of the skin (i.e., wounds, rashes, prickly heat disorder, psoriasis, dermatitis), atherosclerosis, uticaria, eczema, photosensitivity, autoimmune disorders (i.e. lupus erythematosus, vitiligo, dermatomyositis, morphea, scleroderma, 35 pemphigoid, and pemphigus), keloids, striae, erythema, petechiae, purpura, and xanthelasma. Moreover, such disorders may predispose increased susceptibility to viral

and bacterial infections of the skin (i.e. cold sores, warts, chickenpox, molluscum contagiosum, herpes zoster, boils, cellulitis, erysipelas, impetigo, tinea, athletes foot, and ringworm). Alternatively, the tissue distribution within dendritic cells and fetal brain indicates that the protein product of this gene is useful for the detection/treatment of neurodegenerative disease states and behavioral disorders such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses , autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:57 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1155 of SEQ ID NO:57, b is an integer of 15 to 1169, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:57, and where the b is greater than or equal to a + 14.

## 25 FEATURES OF PROTEIN ENCODED BY GENE NO: 48

This gene maps to chromosome 1, and therefore, may be used as a marker in linkage analysis for chromosome 1.

This gene is expressed primarily in ovary tumors and to a lesser extent in breast, placenta, pineal gland, infant brain, T cell and B cell lymphoma.

30 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, ovary cancer, reproductive disorders, particularly of the female reproductive system, as well as disorders of the immune system, including lymphoma and immunodeficiencies. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential

identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system and female reproductive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., reproductive, endocrine tissue, cells and tissue of the immune system, and cancerous and wounded tissues) or bodily fluids (e.g., amniotic fluid, lymph, breast milk, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:130 as residues: Asn-33 to Lys-38.

The tissue distribution in tumors of the ovary and lymph nodes, as well as, in breast and placental tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and intervention of such tumors, in addition to other tumors where expression has been indicated. Protein, as well as, antibodies directed against the protein may show utility as a tissue-specific marker and/or immunotherapy target for the above listed tissues. Alternatively, The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Expression of this gene product in T-cells indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tumors and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:58 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present

invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1052 of SEQ ID NO:58, b is an integer of 15 to 1066, where both a and b 5 correspond to the positions of nucleotide residues shown in SEQ ID NO:58, and where the b is greater than or equal to a + 14.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 49

This gene is expressed primarily in breast tissue.

10 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, breast cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential 15 identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., breast tissue and cancerous and wounded tissues) or bodily fluids (e.g., breast milk, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or 20 cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:131 as residues: Ala-40 to Trp-45.

The tissue distribution in breast tissue indicates that polynucleotides and 25 polypeptides corresponding to this gene are useful for diagnosis and intervention of tumors within this tissue. Alternatively, the expression in breast may suggest that the protein product of this gene is useful for the diagnosis, treatment, and/or prevention of various reproductive system disorders, particularly of the female reproductive system. Protein, as well as, antibodies directed against the protein may show utility as a tissue-specific marker and/or immunotherapy target for the above listed tissues. Many 30 polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:59 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the 35 present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides

comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 758 of SEQ ID NO:59, b is an integer of 15 to 772, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:59, and where the b is greater than or equal to a + 14.

5

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 50

This gene is expressed primarily in fetal liver, retina and to a lesser extent in thyroid and pineal gland.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, hepatic and endocrine disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hematopoiesis and immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., liver, endocrine tissue, and cancerous and wounded tissues) or bodily fluids (e.g., bile, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:132 as residues: Lys-9 to Trp-14.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection and treatment of liver disorders and cancers (e.g. hepatoblastoma, jaundice, hepatitis, liver metabolic diseases and conditions that are attributable to the differentiation of hepatocyte progenitor cells). In addition the expression in fetus would suggest a useful role for the protein product in developmental abnormalities, fetal deficiencies, pre-natal disorders and various wound-healing models and/or tissue trauma. Alternatively, expression in pineal and thyroid glands indicates that the protein product of this gene is useful for the detection, treatment, and/or prevention of various endocrine disorders and cancers, particularly Addison's disease, Cushing's Syndrome, and disorders and/or cancers of the pancreas (e.g. diabetes mellitus), adrenal cortex, ovaries, pituitary (e.g., hyper-, hypopituitarism), thyroid (e.g. hyper-, hypothyroidism), parathyroid (e.g. hyper-, hypoparathyroidism), hypothalamus, and testes. Protein, as well as, antibodies

directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:60 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1184 of SEQ ID NO:60, b is an integer of 15 to 1198, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:60, and where the b is greater than or equal to a + 14.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 51

The translation product of this gene shares sequence homology with seizure-related gene product 6 type 2 precursor which is a novel, brain-specific, gene thought to be important in the predisposition of seizures (See Genebank Accession No.gnllPIDld1006729). One embodiment of this gene comprises polypeptides of the following amino acid sequence: AGIQHELACDNPGLPENGYQILYKRLYLPGESLT FMCYEGFELMGEVTIRCILGQPSHWNGPLPVCKVAE AAAETSLEGGN @ (SEQ ID NO:170) QPSHWNGPLPVCKVAEAAAETSLEGGN @ (SEQ ID NO:171), and/or YETGETREYEVSI (SEQ ID NO:172). An additional embodiment is the polynucleotides encoding these polypeptides.

This gene is expressed primarily in brain. Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, seizures and other neural and CNS disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the brain, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., neural tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an

individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:133 as residues: Cys-26 to Leu-32, Thr-49 to Ile-55, Glu-57 to Glu-63.

The tissue distribution in brain combined with its homology to a putative seizure gene indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of seizures and epilepsy, including neurodegenerative disease states and behavioral disorders such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses , autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:61 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 544 of SEQ ID NO:61, b is an integer of 15 to 558, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:61, and where the b is greater than or equal to a + 14.

## FEATURES OF PROTEIN ENCODED BY GENE NO: 52

When tested against PC12 cell lines, supernatants removed from cells containing this gene activated the EGR1 (early growth response 1) pathway. Thus, it is likely that this gene activates sensory neuron cells through the EGR1 signal transduction pathway. EGR1 is a separate signal transduction pathway from Jak-STAT, genes containing the EGR1 promoter are induced in various tissues and cell types upon activation, leading the cells to undergo differentiation and proliferation.

This gene is expressed primarily in fetal brain.  
Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a

biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurological disorders, particularly of the developing embryo. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For 5 a number of disorders of the above tissues or cells, particularly of the fetal brain and CNS, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or 10 bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:134 as residues: Arg-16 to Thr-35.

The tissue distribution in fetal brain combined with the detected biological activity within sensory neurons indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative 15 disease states and behavioral disorders such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, the gene or gene product may also play a 20 role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many 25 polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:62 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides 30 comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 602 of SEQ ID NO:62, b is an integer of 15 to 616, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:62, and where the b is greater than or equal to a + 14.

35 **FEATURES OF PROTEIN ENCODED BY GENE NO: 53**

This gene is expressed primarily in frontal cortex, and schizophrenoid brain tissue.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, Schizophrenia, and other neurodegenerative disorders, including cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the brain, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., neural tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:135 as residues: Asp-65 to Asn-72.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states and behavioral disorders such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses , autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:63 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 797 of SEQ ID NO:63, b is an integer of 15 to 811, where

both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:63, and where the b is greater than or equal to a + 14.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 54

5 This gene is expressed primarily in osteoclastoma, gall bladder, and infant brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are 10 not limited to, Osteoclastoma, and gastrointestinal, skeletal, and neural disorders, particularly cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the bone, expression of this gene at significantly higher 15 or lower levels may be routinely detected in certain tissues or cell types (e.g., skeletal tissue, neural tissue, and cancerous and wounded tissues) or bodily fluids (e.g., bile, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an 20 individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:136 as residues: Gly-23 to His-30:

The tissue distribution in osteoclastoma cells indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of various skeletal system disorders, particularly bone cancer. Moreover, the 25 tissue distribution within the gall bladder indicates that the protein product of this gene is useful for the detection and treatment of liver disorders and cancers (e.g. hepatoblastoma, jaundice, hepatitis, liver metabolic diseases and conditions that are attributable to the differentiation of hepatocyte progenitor cells). In addition the expression in fetus would suggest a useful role for the protein product in developmental 30 abnormalities, fetal deficiencies, pre-natal disorders and various wound-healing models and/or tissue trauma. Osteoclastoma, Gall Bladder tumor. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:64 and may have been publicly available prior to conception of the present invention. Preferably, 35 such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably

excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 979 of SEQ ID NO:64, b is an integer of 15 to 993, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:64, and where  
5 the b is greater than or equal to a + 14.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 55

The translation product of this gene shares sequence homology with b cell growth factor which is thought to be important in B cell development.

10 This gene is expressed primarily in breast lymph node and primary dendritic cells

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are  
15 not limited to, immune disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell  
20 types (e.g., cells and tissue of the immune system, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include  
25 those comprising a sequence shown in SEQ ID NO:137 as residues: Ser-39 to Tyr-46.

The tissue distribution combined with its homology to a B-cell growth factor indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Expression of this gene product in lymph nodes indicates a role in the regulation of the proliferation;  
30 survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved  
35 in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS,

leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tumors and tissues. In addition, this gene product may have commercial utility in the expansion of stem 5 cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences 10 are related to SEQ ID NO:65 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general 15 formula of a-b, where a is any integer between 1 to 675 of SEQ ID NO:65, b is an integer of 15 to 689, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:65, and where the b is greater than or equal to a + 14.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 56

20 The translation product of this gene shares sequence homology with small hepatocellular oncoprotein which is thought to be important in liver development (See Genebank Accession No. R07057). This gene maps to chromosome 20, and therefore, may be used as a marker in linkage analysis for chromosome 20.

This gene is expressed primarily in monocytes.

25 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, liver disorders, particularly hepatoma. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes 30 for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and digestive systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., haematopoietic cells and tissue, liver, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial 35 fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level

in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:138 as residues: Met-17 to Lys-32.

The homology to a hepatocellular oncogene indicates that polynucleotides and 5 polypeptides corresponding to this gene are useful for the detection and treatment of liver disorders and cancers (e.g. hepatoblastoma, jaundice, hepatitis, liver metabolic diseases and conditions that are attributable to the differentiation of hepatocyte progenitor cells). The expression also indicates a useful role for the protein product in developmental abnormalities, fetal deficiencies, pre-natal disorders and various would-10 healing models and/or tissue trauma. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Alternatively, The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are . useful for the treatment and diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia, 15 thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. The uses include bone marrow cell ex vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia. The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, 20 immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:66 and may 25 have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer 30 between 1 to 928 of SEQ ID NO:66, b is an integer of 15 to 942, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:66, and where the b is greater than or equal to a + 14.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 57

35 The translation product of this gene was shown to have homology to the human proteins myotubularin related protein 3 and NTII-1 nerve protein (See Genebank

Accession Nos. gil1378120 and R99800, respectively) which are both thought to play important roles as growth factors in muscle and nerve tissue, respectively. Preferred polypeptides comprise the following amino acid sequence:

DDDGLPFPTDVIQHRLRQIEAGYKQEVEQLRR

5 QVRDSDEXGHPSLLCPSSRAPMDYEDDFTCLKESDGSDTEFGSDHSEDCLSEA  
SWEPVDKKETEVTRWVPDHMASHCYNCDCEFWLAKRRHHCRNCGNVFCAG  
CCHLKLPIPDQQLYDPVLVCNSCYXTHSSLSCQGTHEPTAEETHCYSFQLNAGE  
KPVQF (SEQ ID NO:173), SEASWEPPVDKKETEVTRWVPDHMASHC (SEQ ID  
NO:174), HHCRNCGNVF (SEQ ID NO:175, and/or RLRQIEAGYKQEVE (SEQ ID  
10 NO:176). Also preferred are the polynucleotides encoding these polypeptides.

This gene is expressed primarily in bone, spleen, brain, apoptotic T cells, hypothalamus, and other immune cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a 15 biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune system, musculoskeletal, and neural disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous 20 and immune systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., cells and tissue of the immune system, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene 25 expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:139 as residues: Glu-63 to Asp-68.

The tissue distribution in neural tissue combined with the homology to a nerve growth factor indicates that polynucleotides and polypeptides corresponding to this 30 gene are useful for the detection/treatment of neurodegenerative disease states and behavioral disorders such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses , autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and 35 perception, as well as neuromuscular disorders such as MS and muscular dystrophy. In addition, the gene or gene product may also play a role in the treatment and/or detection

of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:67 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2295 of SEQ ID NO:67, b is an integer of 15 to 2309, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:67, and where the b is greater than or equal to a + 14.

15

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 58

This gene is expressed primarily in ovary.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, ovarian cancer, and other disorders that afflict the female reproductive system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the female reproductive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., reproductive tissue, and cancerous and wounded tissues) or bodily fluids (e.g., amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:140 as residues: Pro-23 to Gly-54.

The tissue distribution in tumors the ovary indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and intervention of these tumors, in addition to other tumors where expression has been indicated. Protein, as well as, antibodies directed against the protein may show utility as a tissue-specific

marker and/or immunotherapy target for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:68 and may have been publicly available prior to conception of the present invention. Preferably, 5 such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 800 of SEQ ID NO:68, b is an integer of 15 to 814, where both a and b 10 correspond to the positions of nucleotide residues shown in SEQ ID NO:68, and where the b is greater than or equal to a + 14.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 59

This gene is expressed primarily in ovary.

15 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, ovarian cancer, and other disorders afflicting the female reproductive system. Similarly, polypeptides and antibodies directed to these polypeptides are useful 20 in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., reproductive tissue, and cancerous and wounded tissues) or bodily fluids (e.g., amniotic fluid, serum, plasma, 25 urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in ovarian tissue indicates that polynucleotides and 30 polypeptides corresponding to this gene are useful for diagnosis and intervention of ovarian tumors, in addition to other tumors of the female reproductive system. Protein, as well as, antibodies directed against the protein may show utility as a tissue-specific marker and/or immunotherapy target for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through 35 sequence databases. Some of these sequences are related to SEQ ID NO:69 and may have been publicly available prior to conception of the present invention. Preferably,

such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer 5 between 1 to 774 of SEQ ID NO:69, b is an integer of 15 to 788, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:69, and where the b is greater than or equal to a + 14.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 60

When tested against PC12 cell lines, supernatants removed from cells containing this gene activated the EGR1 (early growth response 1) pathway. Thus, it is likely that this gene activates sensory neuronal cells through the EGR1 signal transduction pathway. EGR1 is a separate signal transduction pathway from Jak-STAT, genes containing the EGR1 promoter are induced in various tissues and cell types upon activation, leading the cells to undergo differentiation and proliferation.

This gene is expressed primarily in osteoblast.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, skeletal disorders, particular bone cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the Bone, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., skeletal tissue, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:142 as residues: Tyr-20 to Lys-31.

In addition the expression of this gene product in synovium would suggest a role in the detection and treatment of disorders and conditions affecting the skeletal system, in particular osteoporosis as well as disorders afflicting connective tissues (e.g. arthritis, trauma, tendonitis, chondromalacia and inflammation) as well as in the diagnosis or treatment of various autoimmune disorders such as rheumatoid arthritis, lupus, scleroderma, and dermatomyositis as well as dwarfism, spinal deformation,

and specific joint abnormalities as well as chondrodysplasias i.e. spondyloepiphyseal dysplasia congenita, familial osteoarthritis, Atelosteogenesis type II, metaphyseal chondrodysplasia type Schmid. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:70 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.

Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 777 of SEQ ID NO:70, b is an integer of 15 to 791, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:70, and where the b is greater than or equal to a + 14.

15

## FEATURES OF PROTEIN ENCODED BY GENE NO: 61

This gene is expressed primarily in adipocyte.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, obesity. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., adipose tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:143 as residues: His-2 to Leu-8.

The tissue distribution would suggest that polynucleotides and polypeptides corresponding to this gene are useful in the diagnosis, treatment, and/or prevention of obesity and lipid metabolism disorders. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are

35

publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:71 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is 5 cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 790 of SEQ ID NO:71, b is an integer of 15 to 804, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:71, and where the b is greater than or equal to a + 14.

10

## FEATURES OF PROTEIN ENCODED BY GENE NO: 62

When tested against Jurkat T-cell lines, supernatants removed from cells containing this gene activated the GAS (gamma activation site) pathway. Thus, it is likely that this gene activates T-cells through the Jak-STAT signal transduction 15 pathway. GAS is a promoter element found upstream in many genes which are involved in the Jaks-STAT pathway. The Jaks-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells.

This gene is expressed primarily in synovial tissue. Therefore, polynucleotides and polypeptides of the invention are useful as 20 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, chronic synovitis, immune. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for 25 differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., cells and tissue of the immune system, and skeletal tissue, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having 30 such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:144 as residues: Pro-74 to Lys-82.

The biological assay results indicating activity in Jurkat T-cells for the protein 35 product of this gene indicates that polynucleotides and polypeptides corresponding to this gene useful for the diagnosis and treatment of a variety of immune system

disorders. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene has been shown to activate genes in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tumors and tissues.

5 In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. In addition the expression of this gene product in synovial tissues indicates a role for this protein in the detection and treatment of disorders and conditions affecting the skeletal system, in particular the connective

10 15 tissues (e.g. arthritis, trauma, tendonitis, chondromalacia and inflammation) as well as in the diagnosis or treatment of various autoimmune disorders such as rheumatoid arthritis, lupus, scleroderma, and dermatomyositis as well as dwarfism, spinal deformation, and specific joint abnormalities as well as chondrodysplasias i.e. spondyloepiphyseal dysplasia congenita, familial osteoarthritis, Atelosteogenesis type

20 25 II, metaphyseal chondrodysplasia type Schmid. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:72 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 769 of SEQ ID NO:72, b is an integer of 15 to 783, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:72, and where the b is greater than or equal to a + 14.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 63

The translation product of this gene shares sequence homology with fetal troponin which is known to be essential for normal muscular function in all skeletal and cardiovascular muscles.

This gene is expressed primarily in melanocytes, fetal liver, brain, testes, spleen, and placenta.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, melanoma, neuromuscular disorders, such as multiple sclerosis, and endothelial-related diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., muscle, endothelial cells and tissue, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:145 as residues: Ala-27 to Leu-36, Phe-40 to Leu-50.

The tissue distribution and homology to troponin indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment, diagnosis, and/or prevention of various neuromuscular disorders such as MS, muscular dystrophy, cardiomyopathy, myositis, myomas, leiomyomas, rhabdomyosarcomas, and coronary heart disease. Alternatively, the expression in a variety of fetal immune tissues indicates that the protein product of this gene is useful for the diagnosis and treatment of a variety of immune system disorders. Expression of this gene product in spleen and fetal liver indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tumors and tissues. In addition, this

gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

- 5 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:73 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.
- 10 Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1509 of SEQ ID NO:73, b is an integer of 15 to 1523, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:73, and where the b is greater than or equal to a + 14.

15

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 64

The translation product of this gene shares sequence homology with small hepatocellular oncoprotein gene which is thought to be important in liver disorders (See Genebank Accession No. R07057). One embodiment of this gene comprises polypeptides of the following amino acid sequence:

MSHCARPLFFETFFILLSPRLKCSGTNTVHYSLDLLGSSNSASVPQVGGLTNAQ HDTWLIFVFCVCVCEPLRRPWAFLISVTSSIK (SEQ ID NO:177), and/or VPQVGGLTNAQHDTWLIFVFCVCVCEPLRR (SEQ ID NO:178). An additional embodiment is the polynucleotides encoding these polypeptides.

- 25 This gene is expressed primarily in neutrophils, hemangiopericytoma, activated T cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune and hepatic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hepatoma and immune system disorder, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., cells and tissue of the immune system, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine,

synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:146 as residues: Arg-20 to Gly-25.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Expression of this gene product in neutrophils and T-cells indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions.

Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tumors and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Alternatively, the homology to a hepatic oncogene indicates that the protein product of this gene is useful for the detection and treatment of liver disorders and cancers (e.g. hepatoblastoma, jaundice, hepatitis, liver metabolic diseases and conditions that are attributable to the differentiation of hepatocyte progenitor cells). In addition, the expression would suggest a useful role for the protein product in developmental abnormalities, fetal deficiencies, pre-natal disorders and various wound-healing models and/or tissue trauma. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:74 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.

Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of

a-b, where a is any integer between 1 to 744 of SEQ ID NO:74, b is an integer of 15 to 758, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:74, and where the b is greater than or equal to a + 14.



Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	NT SEQ ID NO: X	5' NT of Total NT Seq.	3' NT of Clone Seq.	5' NT of Start Codon	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
14	HHEPT60	209138 07/03/97	pCMV Sport 3.0	24	532	21	532	245	96	1	18
15	HLWBZ73	209138 07/03/97	pCMV Sport 3.0	25	920	352	905	434	97	1	28
16	HNGFR75	209138 07/03/97	Uni-ZAP XR	26	917	1	917	54	98	1	23
17	HNHEO73	209138 07/03/97	Uni-ZAP XR	27	662	1	662	84	99	1	28
18	HNHFO29	209138 07/03/97	Uni-ZAP XR	28	699	1	699	160	160	100	1
19	HONAH29	209138 07/03/97	pBluescript SK+	29	1637	17	1632	144	144	101	1
20	HGCAB62	209138 07/03/97	pSportI	30	2142	1223	2109	110	110	102	1
20	HGCAB62	209138 07/03/97	pSportI	75	1096	209	1096	406	406	147	1
21	HAQBI01	209138 07/03/97	Uni-ZAP XR	31	1564	1084	1564	231	231	103	1
21	HAQBI01	209138 07/03/97	Uni-ZAP XR	76	1230	406	1171	637	148	1	20
22	HCMSW93	209139 07/03/97	Uni-ZAP XR	32	1631	254	855	362	362	104	1
23	HDPPBA48	209139 07/03/97	pCMV Sport 3.0	33	978	1	978	478	478	105	1
23	HDPPBA48	209139 07/03/97	pCMV Sport 3.0	77	911	1	911	309	149	1	12



Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	NT SEQ	5' NT of Total NT NO:	3' NT of Clone Seq.	5' NT of Start Seq.	First AA of Signal Seq.	5' NT of Codon	AA SEQ ID NO:	First AA AA of Sig Pep	Last AA of Sig Pep	First AA AA of Sig Pep	Last AA of Sig Pep	First AA AA of Sig Pep	Last AA of Sig Pep
37	HPEBT96	209139 07/03/97	Uni-ZAP XR	47	1849	812	1517	968	968	119	1	37	38	45	
38	HPFCE63	209139 07/03/97	Uni-ZAP XR	48	926	203	926	414	414	120	1	17	18	28	
39	HPMBZ15	209139 07/03/97	Uni-ZAP XR	49	1593	182	1593	1124	1124	121	1	49	50	62	
39	HPMBZ15	209139 07/03/97	Uni-ZAP XR	78	488	20	488	121	121	150	1			15	
40	HROAE16	209139 07/03/97	Uni-ZAP XR	50	978	1	978	209	209	122	1	32	33	86	
41	HSAYM40	209139 07/03/97	Uni-ZAP XR	51	433	1	433	190	190	123	1	19	20	63	
42	HSNAD72	209139 07/03/97	Uni-ZAP XR	52	861	1	861	220	220	124	1	22	23	35	
43	HSNAU78	209139 07/03/97	Uni-ZAP XR	53	510	1	510	89	89	125	1	17	18	31	
44	HTBAB28	209139 07/03/97	Uni-ZAP XR	54	309	1	309	43	43	126	1	35	36	40	
45	HALSQ38	209141 07/09/97	Uni-ZAP XR	55	1585	741	1585	776	776	127	1	17	18	27	
46	HAQBT52	209141 07/09/97	Uni-ZAP XR	56	874	20	874	136	136	128	1	47	48	68	
47	HBIBL04	209141 07/09/97	Uni-ZAP XR	57	1169	5	753	726	726	129	1	40	41	87	
47	HBIBL04	209141 07/09/97	Uni-ZAP XR	79	753	1	749	228	228	151	1	21	22	21	

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	NT SEQ ID	NT NO: X	5' NT of Total Seq. Uni-ZAP XR	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of Signal Pep	5' NT of First AA	AA SEQ ID	First AA NO: Y	Last AA Sig Pep	First AA Sig Pep	Last AA Secreted Portion	Last AA of ORF
48	HBJC195	209141 07/09/97	209141 07/09/97	Uni-ZAP XR	58	1066	1	1066	188	188	130	1	12	13	63
49	HBNBQ61	209141 07/09/97	209141 07/09/97	Uni-ZAP XR	59	772	1	772	88	88	131	1			36
50	HE2ID06	209141 07/09/97	209141 07/09/97	Uni-ZAP XR	60	1198	554	1198	715	715	132	1	28	29	40
51	HEBCM63	209141 07/09/97	209141 07/09/97	Uni-ZAP XR	61	558	1	558	246	246	133	1	26	27	68
52	HFFAK76	209141 07/09/97	209141 07/09/97	Lambda ZAP II	62	616	1	616	150	150	134	1	17	18	47
53	HFRBF28	209141 07/09/97	209141 07/09/97	Uni-ZAP XR	63	811	1	811	57	57	135	1	16	17	16
54	HGBHBM89	209141 07/09/97	209141 07/09/97	Uni-ZAP XR	64	993	1	970	49	49	136	1	28	29	54
55	HLMBP18	209141 07/09/97	209141 07/09/97	Lambda ZAP II	65	689	1	689		398	137	1			14
56	HMSEL55	209141 07/09/97	209141 07/09/97	Uni-ZAP XR	66	942	1	942	437	437	138	1	18	19	35
57	HMWCF89	209141 07/09/97	209141 07/09/97	Uni-Zap XR	67	2309	1609	2308	1216	1216	139	1	23	24	80
58	HODAZ55	209141 07/09/97	209141 07/09/97	Uni-ZAP XR	68	814	1	814	37	37	140	1	16	17	24
59	HODDF08	209141 07/09/97	209141 07/09/97	Uni-ZAP XR	69	788	1	788	141	141	141	1	32	33	50

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	NT SEQ ID NO:	NT of Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Seq.	AA SEQ ID NO:	First AA of Signal Pep	Last AA of Signal Pep	First AA of Secreted Pep	Last AA of Secreted Pep	First AA of Portion Y	Last AA of Portion Y	Last AA of ORF
60	HOSDK95	209141 07/09/97	Uni-ZAP XR	70	791	1	791	121	121	142	1	28	29	29	29
61	HOUAR65	209141 07/09/97	Uni-ZAP XR	71	804	1	804		53	143	1	28	29	29	49
62	HSVAC77	209141 07/09/97	Uni-ZAP XR	72	783	1	783	27	27	144	1	28	29	29	86
63	HJAAAT30	209141 07/09/97	pBluescript SK-	73	1523	767	1448	855	855	145	1	23	24	24	57
63	HJAAAT30	209141 07/09/97	pBluescript SK-	81	1327	71	750	386	386	153	1	39	40	40	40
64	HNECF34	209141 07/09/97	Uni-ZAP XR	74	758	1	758	106	106	146	1	23	24	24	87
64	HNECF34	209141 07/09/97	Uni-ZAP XR	82	758	1	758	269	269	154	1	17	18	18	24

Table 1 summarizes the information corresponding to each "Gene No." described above. The nucleotide sequence identified as "NT SEQ ID NO:X" was assembled from partially homologous ("overlapping") sequences obtained from the "cDNA clone ID" identified in Table 1 and, in some cases, from additional related DNA clones. The 5 overlapping sequences were assembled into a single contiguous sequence of high redundancy (usually three to five overlapping sequences at each nucleotide position), resulting in a final sequence identified as SEQ ID NO:X.

The cDNA Clone ID was deposited on the date and given the corresponding deposit number listed in "ATCC Deposit No:Z and Date." Some of the deposits contain 10 multiple different clones corresponding to the same gene. "Vector" refers to the type of vector contained in the cDNA Clone ID.

"Total NT Seq." refers to the total number of nucleotides in the contig identified by "Gene No." The deposited clone may contain all or most of these sequences, reflected by the nucleotide position indicated as "5' NT of Clone Seq." and the "3' NT 15 of Clone Seq." of SEQ ID NO:X. The nucleotide position of SEQ ID NO:X of the putative start codon (methionine) is identified as "5' NT of Start Codon." Similarly, the nucleotide position of SEQ ID NO:X of the predicted signal sequence is identified as "5' NT of First AA of Signal Pep."

The translated amino acid sequence, beginning with the methionine, is identified 20 as "AA SEQ ID NO:Y," although other reading frames can also be easily translated using known molecular biology techniques. The polypeptides produced by these alternative open reading frames are specifically contemplated by the present invention.

The first and last amino acid position of SEQ ID NO:Y of the predicted signal peptide is identified as "First AA of Sig Pep" and "Last AA of Sig Pep." The predicted 25 first amino acid position of SEQ ID NO:Y of the secreted portion is identified as "Predicted First AA of Secreted Portion." Finally, the amino acid position of SEQ ID NO:Y of the last amino acid in the open reading frame is identified as "Last AA of ORF."

SEQ ID NO:X and the translated SEQ ID NO:Y are sufficiently accurate and 30 otherwise suitable for a variety of uses well known in the art and described further below. For instance, SEQ ID NO:X is useful for designing nucleic acid hybridization probes that will detect nucleic acid sequences contained in SEQ ID NO:X or the cDNA contained in the deposited clone. These probes will also hybridize to nucleic acid molecules in biological samples, thereby enabling a variety of forensic and diagnostic 35 methods of the invention. Similarly, polypeptides identified from SEQ ID NO:Y may be used to generate antibodies which bind specifically to the secreted proteins encoded by the cDNA clones identified in Table 1.

Nevertheless, DNA sequences generated by sequencing reactions can contain sequencing errors. The errors exist as misidentified nucleotides, or as insertions or deletions of nucleotides in the generated DNA sequence. The erroneously inserted or deleted nucleotides cause frame shifts in the reading frames of the predicted amino acid sequence. In these cases, the predicted amino acid sequence diverges from the actual amino acid sequence, even though the generated DNA sequence may be greater than 99.9% identical to the actual DNA sequence (for example, one base insertion or deletion in an open reading frame of over 1000 bases).

Accordingly, for those applications requiring precision in the nucleotide sequence or the amino acid sequence, the present invention provides not only the generated nucleotide sequence identified as SEQ ID NO:X and the predicted translated amino acid sequence identified as SEQ ID NO:Y, but also a sample of plasmid DNA containing a human cDNA of the invention deposited with the ATCC, as set forth in Table 1. The nucleotide sequence of each deposited clone can readily be determined by sequencing the deposited clone in accordance with known methods. The predicted amino acid sequence can then be verified from such deposits. Moreover, the amino acid sequence of the protein encoded by a particular clone can also be directly determined by peptide sequencing or by expressing the protein in a suitable host cell containing the deposited human cDNA, collecting the protein, and determining its sequence.

The present invention also relates to the genes corresponding to SEQ ID NO:X, SEQ ID NO:Y, or the deposited clone. The corresponding gene can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include preparing probes or primers from the disclosed sequence and identifying or amplifying the corresponding gene from appropriate sources of genomic material.

Also provided in the present invention are species homologs. Species homologs may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source for the desired homologue.

The polypeptides of the invention can be prepared in any suitable manner. Such polypeptides include isolated naturally occurring polypeptides, recombinantly produced polypeptides, synthetically produced polypeptides, or polypeptides produced by a combination of these methods. Means for preparing such polypeptides are well understood in the art.

The polypeptides may be in the form of the secreted protein, including the mature form, or may be a part of a larger protein, such as a fusion protein (see below).

It is often advantageous to include an additional amino acid sequence which contains secretory or leader sequences, pro-sequences, sequences which aid in purification, such as multiple histidine residues, or an additional sequence for stability during recombinant production.

5        The polypeptides of the present invention are preferably provided in an isolated form, and preferably are substantially purified. A recombinantly produced version of a polypeptide, including the secreted polypeptide, can be substantially purified by the one-step method described in Smith and Johnson, Gene 67:31-40 (1988).  
10      Polypeptides of the invention also can be purified from natural or recombinant sources using antibodies of the invention raised against the secreted protein in methods which are well known in the art.

### Signal Sequences

15      Methods for predicting whether a protein has a signal sequence, as well as the cleavage point for that sequence, are available. For instance, the method of McGeoch, Virus Res. 3:271-286 (1985), uses the information from a short N-terminal charged region and a subsequent uncharged region of the complete (uncleaved) protein. The method of von Heinje, Nucleic Acids Res. 14:4683-4690 (1986) uses the information from the residues surrounding the cleavage site, typically residues -13 to +2, where +1  
20     indicates the amino terminus of the secreted protein. The accuracy of predicting the cleavage points of known mammalian secretory proteins for each of these methods is in the range of 75-80%. (von Heinje, *supra*.) However, the two methods do not always produce the same predicted cleavage point(s) for a given protein.

25      In the present case, the deduced amino acid sequence of the secreted polypeptide was analyzed by a computer program called SignalP (Henrik Nielsen et al., Protein Engineering 10:1-6 (1997)), which predicts the cellular location of a protein based on the amino acid sequence. As part of this computational prediction of localization, the methods of McGeoch and von Heinje are incorporated. The analysis of the amino acid sequences of the secreted proteins described herein by this program provided the results  
30     shown in Table 1.

As one of ordinary skill would appreciate, however, cleavage sites sometimes vary from organism to organism and cannot be predicted with absolute certainty.  
Accordingly, the present invention provides secreted polypeptides having a sequence shown in SEQ ID NO:Y which have an N-terminus beginning within 5 residues (i.e., +  
35     or - 5 residues) of the predicted cleavage point. Similarly, it is also recognized that in some cases, cleavage of the signal sequence from a secreted protein is not entirely

uniform, resulting in more than one secreted species. These polypeptides, and the polynucleotides encoding such polypeptides, are contemplated by the present invention.

Moreover, the signal sequence identified by the above analysis may not necessarily predict the naturally occurring signal sequence. For example, the naturally occurring signal sequence may be further upstream from the predicted signal sequence. However, it is likely that the predicted signal sequence will be capable of directing the secreted protein to the ER. These polypeptides, and the polynucleotides encoding such polypeptides, are contemplated by the present invention.

10 **Polynucleotide and Polypeptide Variants**

"Variant" refers to a polynucleotide or polypeptide differing from the polynucleotide or polypeptide of the present invention, but retaining essential properties thereof. Generally, variants are overall closely similar, and, in many regions, identical to the polynucleotide or polypeptide of the present invention.

15 By a polynucleotide having a nucleotide sequence at least, for example, 95% "identical" to a reference nucleotide sequence of the present invention, it is intended that the nucleotide sequence of the polynucleotide is identical to the reference sequence except that the polynucleotide sequence may include up to five point mutations per each 100 nucleotides of the reference nucleotide sequence encoding the polypeptide. In other 20 words, to obtain a polynucleotide having a nucleotide sequence at least 95% identical to a reference nucleotide sequence, up to 5% of the nucleotides in the reference sequence may be deleted or substituted with another nucleotide, or a number of nucleotides up to 5% of the total nucleotides in the reference sequence may be inserted into the reference sequence. The query sequence may be an entire sequence shown in Table 1, the ORF 25 (open reading frame), or any fragment specified as described herein.

As a practical matter, whether any particular nucleic acid molecule or polypeptide is at least 90%, 95%, 96%, 97%, 98% or 99% identical to a nucleotide sequence of the present invention can be determined conventionally using known computer programs. A preferred method for determining the best overall match between 30 a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag et al. (Comp. App. Biosci. (1990) 6:237-245). In a sequence alignment the query and subject sequences are both DNA sequences. An RNA sequence can be compared by converting U's to T's. The result 35 of said global sequence alignment is in percent identity. Preferred parameters used in a FASTDB alignment of DNA sequences to calculate percent identity are:  
Matrix=Unitary, k-tuple=4, Mismatch Penalty=1, Joining Penalty=30, Randomization

Group Length=0, Cutoff Score=1, Gap Penalty=5, Gap Size Penalty 0.05, Window Size=500 or the lenght of the subject nucleotide sequence, whichever is shorter.

If the subject sequence is shorter than the query sequence because of 5' or 3' deletions, not because of internal deletions, a manual correction must be made to the results. This is becuase the FASTDB program does not account for 5' and 3' truncations of the subject sequence when calculating percent identity. For subject sequences truncated at the 5' or 3' ends, relative to the the query sequence, the percent identity is corrected by calculating the number of bases of the query sequence that are 5' and 3' of the subject sequence, which are not matched/aligned, as a percent of the total bases of the query sequence. Whether a nucleotide is matched/aligned is determined by results of the FASTDB sequence alignment. This percentage is then subtracted from the percent identity, calculated by the above FASTDB program using the specified parameters, to arrive at a final percent identity score. This corrected score is what is used for the purposes of the present invention. Only bases outside the 5' and 3' bases of the subject sequence, as displayed by the FASTDB alignment, which are not matched/aligned with the query sequence, are calculated for the purposes of manually adjusting the percent identity score.

For example, a 90 base subject sequence is aligned to a 100 base query sequence to determine percent identity. The deletions occur at the 5' end of the subject sequence and therefore, the FASTDB alignment does not show a matched/alignement of the first 10 bases at 5' end. The 10 unpaired bases represent 10% of the sequence (number of bases at the 5' and 3' ends not matched/total number of bases in the query sequence) so 10% is subtracted from the percent identity score calculated by the FASTDB program. If the remaining 90 bases were perfectly matched the final percent identity would be 90%. In another example, a 90 base subject sequence is compared with a 100 base query sequence. This time the deletions are internal deletions so that there are no bases on the 5' or 3' of the subject sequence which are not matched/aligned with the query. In this case the percent identity calculated by FASTDB is not manually corrected. Once again, only bases 5' and 3' of the subject sequence which are not matched/aligned with the query sequnce are manually corrected for. No other manual corrections are to made for the purposes of the present invention.

By a polypeptide having an amino acid sequence at least, for example, 95% "identical" to a query amino acid sequence of the present invention, it is intended that the amino acid sequence of the subject polypeptide is identical to the query sequence except that the subject polypeptide sequence may include up to five amino acid alterations per each 100 amino acids of the query amino acid sequence. In other words, to obtain a polypeptide having an amino acid sequence at least 95% identical to a query

amino acid sequence, up to 5% of the amino acid residues in the subject sequence may be inserted, deleted, (indels) or substituted with another amino acid. These alterations of the reference sequence may occur at the amino or carboxy terminal positions of the reference amino acid sequence or anywhere between those terminal positions,

5 interspersed either individually among residues in the reference sequence or in one or more contiguous groups within the reference sequence.

As a practical matter, whether any particular polypeptide is at least 90%, 95%, 96%, 97%, 98% or 99% identical to, for instance, the amino acid sequences shown in Table 1 or to the amino acid sequence encoded by deposited DNA clone can be

10 determined conventionally using known computer programs. A preferred method for determining the best overall match between a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag et al. (Comp. App. Biosci. (1990) 6:237-245). In a sequence alignment the query and

15 subject sequences are either both nucleotide sequences or both amino acid sequences. The result of said global sequence alignment is in percent identity. Preferred parameters used in a FASTDB amino acid alignment are: Matrix=PAM 0, k-tuple=2, Mismatch Penalty=1, Joining Penalty=20, Randomization Group Length=0, Cutoff Score=1, Window Size=sequence length, Gap Penalty=5, Gap Size Penalty=0.05, Window

20 Size=500 or the length of the subject amino acid sequence, whichever is shorter.

If the subject sequence is shorter than the query sequence due to N- or C-terminal deletions, not because of internal deletions, a manual correction must be made to the results. This is because the FASTDB program does not account for N- and C-terminal truncations of the subject sequence when calculating global percent identity.

25 For subject sequences truncated at the N- and C-termini, relative to the query sequence, the percent identity is corrected by calculating the number of residues of the query sequence that are N- and C-terminal of the subject sequence, which are not matched/aligned with a corresponding subject residue, as a percent of the total bases of the query sequence. Whether a residue is matched/aligned is determined by results of

30 the FASTDB sequence alignment. This percentage is then subtracted from the percent identity, calculated by the above FASTDB program using the specified parameters, to arrive at a final percent identity score. This final percent identity score is what is used for the purposes of the present invention. Only residues to the N- and C-termini of the subject sequence, which are not matched/aligned with the query sequence, are

35 considered for the purposes of manually adjusting the percent identity score. That is, only query residue positions outside the farthest N- and C-terminal residues of the subject sequence.

For example, a 90 amino acid residue subject sequence is aligned with a 100 residue query sequence to determine percent identity. The deletion occurs at the N-terminus of the subject sequence and therefore, the FASTDB alignment does not show a matching/alignment of the first 10 residues at the N-terminus. The 10 unpaired 5 residues represent 10% of the sequence (number of residues at the N- and C- termini not matched/total number of residues in the query sequence) so 10% is subtracted from the percent identity score calculated by the FASTDB program. If the remaining 90 residues were perfectly matched the final percent identity would be 90%. In another 10 example, a 90 residue subject sequence is compared with a 100 residue query sequence. This time the deletions are internal deletions so there are no residues at the N- or C- termini of the subject sequence which are not matched/aligned with the query. In this 15 case the percent identity calculated by FASTDB is not manually corrected. Once again, only residue positions outside the N- and C-terminal ends of the subject sequence, as displayed in the FASTDB alignment, which are not matched/aligned with the query sequence are manually corrected for. No other manual corrections are to be made for the purposes of the present invention.

The variants may contain alterations in the coding regions, non-coding regions, or both. Especially preferred are polynucleotide variants containing alterations which produce silent substitutions, additions, or deletions, but do not alter the properties or 20 activities of the encoded polypeptide. Nucleotide variants produced by silent substitutions due to the degeneracy of the genetic code are preferred. Moreover, variants in which 5-10, 1-5, or 1-2 amino acids are substituted, deleted, or added in any combination are also preferred. Polynucleotide variants can be produced for a variety of reasons, e.g., to optimize codon expression for a particular host (change codons in 25 the human mRNA to those preferred by a bacterial host such as *E. coli*).

Naturally occurring variants are called "allelic variants," and refer to one of several alternate forms of a gene occupying a given locus on a chromosome of an organism. (Genes II, Lewin, B., ed., John Wiley & Sons, New York (1985).) These allelic variants can vary at either the polynucleotide and/or polypeptide level. 30 Alternatively, non-naturally occurring variants may be produced by mutagenesis techniques or by direct synthesis.

Using known methods of protein engineering and recombinant DNA technology, variants may be generated to improve or alter the characteristics of the polypeptides of the present invention. For instance, one or more amino acids can be deleted from the N-terminus or C-terminus of the secreted protein without substantial 35 loss of biological function. The authors of Ron et al., J. Biol. Chem. 268: 2984-2988 (1993), reported variant KGF proteins having heparin binding activity even after

deleting 3, 8, or 27 amino-terminal amino acid residues. Similarly, Interferon gamma exhibited up to ten times higher activity after deleting 8-10 amino acid residues from the carboxy terminus of this protein. (Dobeli et al., J. Biotechnology 7:199-216 (1988).)

Moreover, ample evidence demonstrates that variants often retain a biological activity similar to that of the naturally occurring protein. For example, Gayle and coworkers (J. Biol. Chem 268:22105-22111 (1993)) conducted extensive mutational analysis of human cytokine IL-1 $\alpha$ . They used random mutagenesis to generate over 3,500 individual IL-1 $\alpha$  mutants that averaged 2.5 amino acid changes per variant over the entire length of the molecule. Multiple mutations were examined at every possible amino acid position. The investigators found that "[m]ost of the molecule could be altered with little effect on either [binding or biological activity]." (See, Abstract.) In fact, only 23 unique amino acid sequences, out of more than 3,500 nucleotide sequences examined, produced a protein that significantly differed in activity from wild-type.

Furthermore, even if deleting one or more amino acids from the N-terminus or C-terminus of a polypeptide results in modification or loss of one or more biological functions, other biological activities may still be retained. For example, the ability of a deletion variant to induce and/or to bind antibodies which recognize the secreted form will likely be retained when less than the majority of the residues of the secreted form are removed from the N-terminus or C-terminus. Whether a particular polypeptide lacking N- or C-terminal residues of a protein retains such immunogenic activities can readily be determined by routine methods described herein and otherwise known in the art.

Thus, the invention further includes polypeptide variants which show substantial biological activity. Such variants include deletions, insertions, inversions, repeats, and substitutions selected according to general rules known in the art so as have little effect on activity. For example, guidance concerning how to make phenotypically silent amino acid substitutions is provided in Bowie, J. U. et al., Science 247:1306-1310 (1990), wherein the authors indicate that there are two main strategies for studying the tolerance of an amino acid sequence to change.

The first strategy exploits the tolerance of amino acid substitutions by natural selection during the process of evolution. By comparing amino acid sequences in different species, conserved amino acids can be identified. These conserved amino acids are likely important for protein function. In contrast, the amino acid positions where substitutions have been tolerated by natural selection indicates that these positions are not critical for protein function. Thus, positions tolerating amino acid substitution could be modified while still maintaining biological activity of the protein.

The second strategy uses genetic engineering to introduce amino acid changes at specific positions of a cloned gene to identify regions critical for protein function. For example, site directed mutagenesis or alanine-scanning mutagenesis (introduction of single alanine mutations at every residue in the molecule) can be used. (Cunningham 5 and Wells, Science 244:1081-1085 (1989).) The resulting mutant molecules can then be tested for biological activity.

As the authors state, these two strategies have revealed that proteins are surprisingly tolerant of amino acid substitutions. The authors further indicate which amino acid changes are likely to be permissive at certain amino acid positions in the 10 protein. For example, most buried (within the tertiary structure of the protein) amino acid residues require nonpolar side chains, whereas few features of surface side chains are generally conserved. Moreover, tolerated conservative amino acid substitutions involve replacement of the aliphatic or hydrophobic amino acids Ala, Val, Leu and Ile; replacement of the hydroxyl residues Ser and Thr; replacement of the acidic residues 15 Asp and Glu; replacement of the amide residues Asn and Gln, replacement of the basic residues Lys, Arg, and His; replacement of the aromatic residues Phe, Tyr, and Trp, and replacement of the small-sized amino acids Ala, Ser, Thr, Met, and Gly.

Besides conservative amino acid substitution, variants of the present invention include (i) substitutions with one or more of the non-conserved amino acid residues, 20 where the substituted amino acid residues may or may not be one encoded by the genetic code, or (ii) substitution with one or more of amino acid residues having a substituent group, or (iii) fusion of the mature polypeptide with another compound, such as a compound to increase the stability and/or solubility of the polypeptide (for example, polyethylene glycol), or (iv) fusion of the polypeptide with additional amino 25 acids, such as an IgG Fc fusion region peptide; or leader or secretory sequence, or a sequence facilitating purification. Such variant polypeptides are deemed to be within the scope of those skilled in the art from the teachings herein.

For example, polypeptide variants containing amino acid substitutions of charged amino acids with other charged or neutral amino acids may produce proteins 30 with improved characteristics, such as less aggregation. Aggregation of pharmaceutical formulations both reduces activity and increases clearance due to the aggregate's immunogenic activity. (Pinckard et al., Clin. Exp. Immunol. 2:331-340 (1967); Robbins et al., Diabetes 36: 838-845 (1987); Cleland et al., Crit. Rev. Therapeutic Drug Carrier Systems 10:307-377 (1993).)

In the present invention, a "polynucleotide fragment" refers to a short polynucleotide having a nucleic acid sequence contained in the deposited clone or shown in SEQ ID NO:X. The short nucleotide fragments are preferably at least about 15 nt, and more preferably at least about 20 nt, still more preferably at least about 30 nt, 5 and even more preferably, at least about 40 nt in length. A fragment "at least 20 nt in length," for example, is intended to include 20 or more contiguous bases from the cDNA sequence contained in the deposited clone or the nucleotide sequence shown in SEQ ID NO:X. These nucleotide fragments are useful as diagnostic probes and primers as discussed herein. Of course, larger fragments (e.g., 50, 150, 500, 600, 2000 10 nucleotides) are preferred.

Moreover, representative examples of polynucleotide fragments of the invention, include, for example, fragments having a sequence from about nucleotide number 1-50, 51-100, 101-150, 151-200, 201-250, 251-300, 301-350, 351-400, 401-450, 451-500, 501-550, 551-600, 651-700, 701-750, 751-800, 800-850, 851-900, 15 901-950, 951-1000, 1001-1050, 1051-1100, 1101-1150, 1151-1200, 1201-1250, 1251-1300, 1301-1350, 1351-1400, 1401-1450, 1451-1500, 1501-1550, 1551-1600, 1601-1650, 1651-1700, 1701-1750, 1751-1800, 1801-1850, 1851-1900, 1901-1950, 20 1951-2000, or 2001 to the end of SEQ ID NO:X or the cDNA contained in the deposited clone. In this context "about" includes the particularly recited ranges, larger or smaller by several (5, 4, 3, 2, or 1) nucleotides, at either terminus or at both termini. Preferably, these fragments encode a polypeptide which has biological activity. More preferably, these polynucleotides can be used as probes or primers as discussed herein.

In the present invention, a "polypeptide fragment" refers to a short amino acid sequence contained in SEQ ID NO:Y or encoded by the cDNA contained in the deposited clone. Protein fragments may be "free-standing," or comprised within a larger polypeptide of which the fragment forms a part or region, most preferably as a single continuous region. Representative examples of polypeptide fragments of the invention, include, for example, fragments from about amino acid number 1-20, 21-40, 41-60, 61-80, 81-100, 102-120, 121-140, 141-160, or 161 to the end of the coding 25 region. Moreover, polypeptide fragments can be about 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, or 150 amino acids in length. In this context "about" includes the particularly recited ranges, larger or smaller by several (5, 4, 3, 2, or 1) amino acids, at either extreme or at both extremes.

Preferred polypeptide fragments include the secreted protein as well as the mature form. Further preferred polypeptide fragments include the secreted protein or the mature form having a continuous series of deleted residues from the amino or the carboxy terminus, or both. For example, any number of amino acids, ranging from 1- 35

60, can be deleted from the amino terminus of either the secreted polypeptide or the mature form. Similarly, any number of amino acids, ranging from 1-30, can be deleted from the carboxy terminus of the secreted protein or mature form. Furthermore, any combination of the above amino and carboxy terminus deletions are preferred.

5      Similarly, polynucleotide fragments encoding these polypeptide fragments are also preferred.

Particularly, N-terminal deletions of the polypeptide of the present invention can be described by the general formula m-p, where p is the total number of amino acids in the polypeptide and m is an integer from 2 to (p-1), and where both of these integers (m & p) correspond to the position of the amino acid residue identified in SEQ ID NO:Y.

Moreover, C-terminal deletions of the polypeptide of the present invention can also be described by the general formula 1-n, where n is an integer from 2 to (p-1), and again where these integers (n & p) correspond to the position of the amino acid residue identified in SEQ ID NO:Y.

15     The invention also provides polypeptides having one or more amino acids deleted from both the amino and the carboxyl termini, which may be described generally as having residues m-n of SEQ ID NO:Y, where m and n are integers as described above.

Also preferred are polypeptide and polynucleotide fragments characterized by structural or functional domains, such as fragments that comprise alpha-helix and alpha-helix forming regions, beta-sheet and beta-sheet-forming regions, turn and turn-forming regions, coil and coil-forming regions, hydrophilic regions, hydrophobic regions, alpha amphipathic regions, beta amphipathic regions, flexible regions, surface-forming regions, substrate binding region, and high antigenic index regions.

20     Polypeptide fragments of SEQ ID NO:Y falling within conserved domains are specifically contemplated by the present invention. Moreover, polynucleotide fragments encoding these domains are also contemplated.

Other preferred fragments are biologically active fragments. Biologically active fragments are those exhibiting activity similar, but not necessarily identical, to an activity of the polypeptide of the present invention. The biological activity of the fragments may include an improved desired activity, or a decreased undesirable activity.

### Epitopes & Antibodies

In the present invention, "epitopes" refer to polypeptide fragments having antigenic or immunogenic activity in an animal, especially in a human. A preferred embodiment of the present invention relates to a polypeptide fragment comprising an epitope, as well as the polynucleotide encoding this fragment. A region of a protein

molecule to which an antibody can bind is defined as an "antigenic epitope." In contrast, an "immunogenic epitope" is defined as a part of a protein that elicits an antibody response. (See, for instance, Geysen et al., Proc. Natl. Acad. Sci. USA 81:3998- 4002 (1983).)

5       Fragments which function as epitopes may be produced by any conventional means. (See, e.g., Houghten, R. A., Proc. Natl. Acad. Sci. USA 82:5131-5135 (1985) further described in U.S. Patent No. 4,631,211.)

In the present invention, antigenic epitopes preferably contain a sequence of at least seven, more preferably at least nine, and most preferably between about 15 to 10 about 30 amino acids. Antigenic epitopes are useful to raise antibodies, including monoclonal antibodies, that specifically bind the epitope. (See, for instance, Wilson et al., Cell 37:767-778 (1984); Sutcliffe, J. G. et al., Science 219:660-666 (1983).)

15      Similarly, immunogenic epitopes can be used to induce antibodies according to methods well known in the art. (See, for instance, Sutcliffe et al., *supra*; Wilson et al., *supra*; Chow, M. et al., Proc. Natl. Acad. Sci. USA 82:910-914; and Bittle, F. J. et al., J. Gen. Virol. 66:2347-2354 (1985).) A preferred immunogenic epitope includes the secreted protein. The immunogenic epitopes may be presented together with a carrier protein, such as an albumin, to an animal system (such as rabbit or mouse) or, if it is long enough (at least about 25 amino acids), without a carrier. However, 20 immunogenic epitopes comprising as few as 8 to 10 amino acids have been shown to be sufficient to raise antibodies capable of binding to, at the very least, linear epitopes in a denatured polypeptide (e.g., in Western blotting.)

As used herein, the term "antibody" (Ab) or "monoclonal antibody" (Mab) is meant to include intact molecules as well as antibody fragments (such as, for example, 25 Fab and F(ab')2 fragments) which are capable of specifically binding to protein. Fab and F(ab')2 fragments lack the Fc fragment of intact antibody, clear more rapidly from the circulation, and may have less non-specific tissue binding than an intact antibody. (Wahl et al., J. Nucl. Med. 24:316-325 (1983).) Thus, these fragments are preferred, as well as the products of a FAB or other immunoglobulin expression library. 30 Moreover, antibodies of the present invention include chimeric, single chain, and humanized antibodies.

### **Fusion Proteins**

Any polypeptide of the present invention can be used to generate fusion proteins. For example, the polypeptide of the present invention, when fused to a 35 second protein, can be used as an antigenic tag. Antibodies raised against the polypeptide of the present invention can be used to indirectly detect the second protein

by binding to the polypeptide. Moreover, because secreted proteins target cellular locations based on trafficking signals, the polypeptides of the present invention can be used as targeting molecules once fused to other proteins.

Examples of domains that can be fused to polypeptides of the present invention 5 include not only heterologous signal sequences, but also other heterologous functional regions. The fusion does not necessarily need to be direct, but may occur through linker sequences.

Moreover, fusion proteins may also be engineered to improve characteristics of 10 the polypeptide of the present invention. For instance, a region of additional amino acids, particularly charged amino acids, may be added to the N-terminus of the polypeptide to improve stability and persistence during purification from the host cell or subsequent handling and storage. Also, peptide moieties may be added to the 15 polypeptide to facilitate purification. Such regions may be removed prior to final preparation of the polypeptide. The addition of peptide moieties to facilitate handling of polypeptides are familiar and routine techniques in the art.

Moreover, polypeptides of the present invention, including fragments, and 20 specifically epitopes, can be combined with parts of the constant domain of immunoglobulins (IgG), resulting in chimeric polypeptides. These fusion proteins facilitate purification and show an increased half-life in vivo. One reported example describes chimeric proteins consisting of the first two domains of the human CD4- 25 polypeptide and various domains of the constant regions of the heavy or light chains of mammalian immunoglobulins. (EP A 394,827; Traunecker et al., Nature 331:84-86 (1988).) Fusion proteins having disulfide-linked dimeric structures (due to the IgG) can also be more efficient in binding and neutralizing other molecules, than the monomeric secreted protein or protein fragment alone. (Fountoulakis et al., J. Biochem. 270:3958-3964 (1995).)

Similarly, EP-A-O 464 533 (Canadian counterpart 2045869) discloses fusion 30 proteins comprising various portions of constant region of immunoglobulin molecules together with another human protein or part thereof. In many cases, the Fc part in a fusion protein is beneficial in therapy and diagnosis, and thus can result in, for example, improved pharmacokinetic properties. (EP-A 0232 262.) Alternatively, deleting the Fc part after the fusion protein has been expressed, detected, and purified, would be desired. For example, the Fc portion may hinder therapy and diagnosis if the fusion protein is used as an antigen for immunizations. In drug discovery, for 35 example, human proteins, such as hIL-5, have been fused with Fc portions for the purpose of high-throughput screening assays to identify antagonists of hIL-5. (See, D.

Bennett et al., J. Molecular Recognition 8:52-58 (1995); K. Johanson et al., J. Biol. Chem. 270:9459-9471 (1995).)

Moreover, the polypeptides of the present invention can be fused to marker sequences, such as a peptide which facilitates purification of the fused polypeptide. In 5 preferred embodiments, the marker amino acid sequence is a hexa-histidine peptide, such as the tag provided in a pQE vector (QIAGEN, Inc., 9259 Eton Avenue, Chatsworth, CA, 91311), among others, many of which are commercially available. As described in Gentz et al., Proc. Natl. Acad. Sci. USA 86:821-824 (1989), for instance, hexa-histidine provides for convenient purification of the fusion protein. 10 Another peptide tag useful for purification, the "HA" tag, corresponds to an epitope derived from the influenza hemagglutinin protein. (Wilson et al., Cell 37:767 (1984).)

Thus, any of these above fusions can be engineered using the polynucleotides or the polypeptides of the present invention.

## 15 Vectors, Host Cells, and Protein Production

The present invention also relates to vectors containing the polynucleotide of the present invention, host cells, and the production of polypeptides by recombinant techniques. The vector may be, for example, a phage, plasmid, viral, or retroviral vector. Retroviral vectors may be replication competent or replication defective. In the 20 latter case, viral propagation generally will occur only in complementing host cells.

The polynucleotides may be joined to a vector containing a selectable marker for propagation in a host. Generally, a plasmid vector is introduced in a precipitate, such as a calcium phosphate precipitate, or in a complex with a charged lipid. If the vector is a virus, it may be packaged in vitro using an appropriate packaging cell line and then 25 transduced into host cells.

The polynucleotide insert should be operatively linked to an appropriate promoter, such as the phage lambda PL promoter, the E. coli lac, trp, phoA and tac promoters, the SV40 early and late promoters and promoters of retroviral LTRs, to name a few. Other suitable promoters will be known to the skilled artisan. The 30 expression constructs will further contain sites for transcription initiation, termination, and, in the transcribed region, a ribosome binding site for translation. The coding portion of the transcripts expressed by the constructs will preferably include a translation initiating codon at the beginning and a termination codon (UAA, UGA or UAG) appropriately positioned at the end of the polypeptide to be translated.

As indicated, the expression vectors will preferably include at least one 35 selectable marker. Such markers include dihydrofolate reductase, G418 or neomycin resistance for eukaryotic cell culture and tetracycline, kanamycin or ampicillin resistance

genes for culturing in *E. coli* and other bacteria. Representative examples of appropriate hosts include, but are not limited to, bacterial cells, such as *E. coli*, *Streptomyces* and *Salmonella typhimurium* cells; fungal cells, such as yeast cells; insect cells such as *Drosophila S2* and *Spodoptera Sf9* cells; animal cells such as CHO, COS, 5 293, and Bowes melanoma cells; and plant cells. Appropriate culture mediums and conditions for the above-described host cells are known in the art.

Among vectors preferred for use in bacteria include pQE70, pQE60 and pQE-9, available from QIAGEN, Inc.; pBluescript vectors, Phagescript vectors, pNH8A, 10 pNH16a, pNH18A, pNH46A, available from Stratagene Cloning Systems, Inc.; and ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 available from Pharmacia Biotech, Inc. Among preferred eukaryotic vectors are pWLNEO, pSV2CAT, pOG44, pXT1 and pSG available from Stratagene; and pSVK3, pBPV, pMSG and pSVL available from Pharmacia. Other suitable vectors will be readily apparent to the skilled artisan.

Introduction of the construct into the host cell can be effected by calcium 15 phosphate transfection, DEAE-dextran mediated transfection, cationic lipid-mediated transfection, electroporation, transduction, infection, or other methods. Such methods are described in many standard laboratory manuals, such as Davis et al., Basic Methods In Molecular Biology (1986). It is specifically contemplated that the polypeptides of the present invention may in fact be expressed by a host cell lacking a recombinant vector.

A polypeptide of this invention can be recovered and purified from recombinant 20 cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography ("HPLC") is employed for purification.

Polypeptides of the present invention, and preferably the secreted form, can also 30 be recovered from: products purified from natural sources, including bodily fluids, tissues and cells, whether directly isolated or cultured; products of chemical synthetic procedures; and products produced by recombinant techniques from a prokaryotic or eukaryotic host, including, for example, bacterial, yeast, higher plant, insect, and mammalian cells. Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated or may be non-glycosylated. In addition, polypeptides of the invention may also include an initial modified methionine residue, in some cases as a result of host-mediated processes. Thus, it is well known in the art that the N-terminal methionine encoded by the 35 translation initiation codon generally is removed with high efficiency from any protein

after translation in all eukaryotic cells. While the N-terminal methionine on most proteins also is efficiently removed in most prokaryotes, for some proteins, this prokaryotic removal process is inefficient, depending on the nature of the amino acid to which the N-terminal methionine is covalently linked.

5

### Uses of the Polynucleotides

Each of the polynucleotides identified herein can be used in numerous ways as reagents. The following description should be considered exemplary and utilizes known techniques.

10        The polynucleotides of the present invention are useful for chromosome identification. There exists an ongoing need to identify new chromosome markers, since few chromosome marking reagents, based on actual sequence data (repeat polymorphisms), are presently available. Each polynucleotide of the present invention can be used as a chromosome marker.

15        Briefly, sequences can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp) from the sequences shown in SEQ ID NO:X. Primers can be selected using computer analysis so that primers do not span more than one predicted exon in the genomic DNA. These primers are then used for PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids containing 20        the human gene corresponding to the SEQ ID NO:X will yield an amplified fragment.

25        Similarly, somatic hybrids provide a rapid method of PCR mapping the polynucleotides to particular chromosomes. Three or more clones can be assigned per day using a single thermal cycler. Moreover, sublocalization of the polynucleotides can be achieved with panels of specific chromosome fragments. Other gene mapping strategies that can be used include *in situ* hybridization, prescreening with labeled flow-sorted chromosomes, and preselection by hybridization to construct chromosome specific-cDNA libraries.

30        Precise chromosomal location of the polynucleotides can also be achieved using fluorescence *in situ* hybridization (FISH) of a metaphase chromosomal spread. This technique uses polynucleotides as short as 500 or 600 bases; however, polynucleotides 2,000-4,000 bp are preferred. For a review of this technique, see Verma et al., "Human Chromosomes: a Manual of Basic Techniques," Pergamon Press, New York (1988).

35        For chromosome mapping, the polynucleotides can be used individually (to mark a single chromosome or a single site on that chromosome) or in panels (for marking multiple sites and/or multiple chromosomes). Preferred polynucleotides correspond to the noncoding regions of the cDNAs because the coding sequences are

more likely conserved within gene families, thus increasing the chance of cross hybridization during chromosomal mapping.

Once a polynucleotide has been mapped to a precise chromosomal location, the physical position of the polynucleotide can be used in linkage analysis. Linkage analysis establishes coinheritance between a chromosomal location and presentation of a particular disease. (Disease mapping data are found, for example, in V. McKusick, Mendelian Inheritance in Man (available on line through Johns Hopkins University Welch Medical Library).) Assuming 1 megabase mapping resolution and one gene per 20 kb, a cDNA precisely localized to a chromosomal region associated with the disease could be one of 50-500 potential causative genes.

Thus, once coinheritance is established, differences in the polynucleotide and the corresponding gene between affected and unaffected individuals can be examined. First, visible structural alterations in the chromosomes, such as deletions or translocations, are examined in chromosome spreads or by PCR. If no structural alterations exist, the presence of point mutations are ascertained. Mutations observed in some or all affected individuals, but not in normal individuals, indicates that the mutation may cause the disease. However, complete sequencing of the polypeptide and the corresponding gene from several normal individuals is required to distinguish the mutation from a polymorphism. If a new polymorphism is identified, this polymorphic polypeptide can be used for further linkage analysis.

Furthermore, increased or decreased expression of the gene in affected individuals as compared to unaffected individuals can be assessed using polynucleotides of the present invention. Any of these alterations (altered expression, chromosomal rearrangement, or mutation) can be used as a diagnostic or prognostic marker.

In addition to the foregoing, a polynucleotide can be used to control gene expression through triple helix formation or antisense DNA or RNA. Both methods rely on binding of the polynucleotide to DNA or RNA. For these techniques, preferred polynucleotides are usually 20 to 40 bases in length and complementary to either the region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 241:456 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Okano, J. Neurochem. 56:560 (1991); Oligodeoxy-nucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988).) Triple helix formation optimally results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques are effective in model

systems, and the information disclosed herein can be used to design antisense or triple helix polynucleotides in an effort to treat disease.

Polynucleotides of the present invention are also useful in gene therapy. One goal of gene therapy is to insert a normal gene into an organism having a defective gene, in an effort to correct the genetic defect. The polynucleotides disclosed in the present invention offer a means of targeting such genetic defects in a highly accurate manner. Another goal is to insert a new gene that was not present in the host genome, thereby producing a new trait in the host cell.

The polynucleotides are also useful for identifying individuals from minute biological samples. The United States military, for example, is considering the use of restriction fragment length polymorphism (RFLP) for identification of its personnel. In this technique, an individual's genomic DNA is digested with one or more restriction enzymes, and probed on a Southern blot to yield unique bands for identifying personnel. This method does not suffer from the current limitations of "Dog Tags" which can be lost, switched, or stolen, making positive identification difficult. The polynucleotides of the present invention can be used as additional DNA markers for RFLP.

The polynucleotides of the present invention can also be used as an alternative to RFLP, by determining the actual base-by-base DNA sequence of selected portions of an individual's genome. These sequences can be used to prepare PCR primers for amplifying and isolating such selected DNA, which can then be sequenced. Using this technique, individuals can be identified because each individual will have a unique set of DNA sequences. Once an unique ID database is established for an individual, positive identification of that individual, living or dead, can be made from extremely small tissue samples.

Forensic biology also benefits from using DNA-based identification techniques as disclosed herein. DNA sequences taken from very small biological samples such as tissues, e.g., hair or skin, or body fluids, e.g., blood, saliva, semen, etc., can be amplified using PCR. In one prior art technique, gene sequences amplified from polymorphic loci, such as DQa class II HLA gene, are used in forensic biology to identify individuals. (Erlich, H., PCR Technology, Freeman and Co. (1992).) Once these specific polymorphic loci are amplified, they are digested with one or more restriction enzymes, yielding an identifying set of bands on a Southern blot probed with DNA corresponding to the DQa class II HLA gene. Similarly, polynucleotides of the present invention can be used as polymorphic markers for forensic purposes.

There is also a need for reagents capable of identifying the source of a particular tissue. Such need arises, for example, in forensics when presented with tissue of

unknown origin. Appropriate reagents can comprise, for example, DNA probes or primers specific to particular tissue prepared from the sequences of the present invention. Panels of such reagents can identify tissue by species and/or by organ type. In a similar fashion, these reagents can be used to screen tissue cultures for contamination.

In the very least, the polynucleotides of the present invention can be used as molecular weight markers on Southern gels, as diagnostic probes for the presence of a specific mRNA in a particular cell type, as a probe to "subtract-out" known sequences in the process of discovering novel polynucleotides, for selecting and making oligomers for attachment to a "gene chip" or other support, to raise anti-DNA antibodies using DNA immunization techniques, and as an antigen to elicit an immune response.

### Uses of the Polypeptides

Each of the polypeptides identified herein can be used in numerous ways. The following description should be considered exemplary and utilizes known techniques.

A polypeptide of the present invention can be used to assay protein levels in a biological sample using antibody-based techniques. For example, protein expression in tissues can be studied with classical immunohistological methods. (Jalkanen, M., et al., J. Cell. Biol. 101:976-985 (1985); Jalkanen, M., et al., J. Cell. Biol. 105:3087-3096 (1987).) Other antibody-based methods useful for detecting protein gene expression include immunoassays, such as the enzyme linked immunosorbent assay (ELISA) and the radioimmunoassay (RIA). Suitable antibody assay labels are known in the art and include enzyme labels, such as, glucose oxidase, and radioisotopes, such as iodine (125I, 121I), carbon (14C), sulfur (35S), tritium (3H), indium (112In), and technetium (99mTc), and fluorescent labels, such as fluorescein and rhodamine, and biotin.

In addition to assaying secreted protein levels in a biological sample, proteins can also be detected *in vivo* by imaging. Antibody labels or markers for *in vivo* imaging of protein include those detectable by X-radiography, NMR or ESR. For X-radiography, suitable labels include radioisotopes such as barium or cesium, which emit detectable radiation but are not overtly harmful to the subject. Suitable markers for NMR and ESR include those with a detectable characteristic spin, such as deuterium, which may be incorporated into the antibody by labeling of nutrients for the relevant hybridoma.

A protein-specific antibody or antibody fragment which has been labeled with an appropriate detectable imaging moiety, such as a radioisotope (for example, 131I, 112In, 99mTc), a radio-opaque substance, or a material detectable by nuclear magnetic

resonance, is introduced (for example, parenterally, subcutaneously, or intraperitoneally) into the mammal. It will be understood in the art that the size of the subject and the imaging system used will determine the quantity of imaging moiety needed to produce diagnostic images. In the case of a radioisotope moiety, for a human 5 subject, the quantity of radioactivity injected will normally range from about 5 to 20 millicuries of 99mTc. The labeled antibody or antibody fragment will then preferentially accumulate at the location of cells which contain the specific protein. In vivo tumor imaging is described in S.W. Burchiel et al., "Immunopharmacokinetics of Radiolabeled Antibodies and Their Fragments." (Chapter 13 in Tumor Imaging: The 10 Radiochemical Detection of Cancer, S.W. Burchiel and B. A. Rhodes, eds., Masson Publishing Inc. (1982).)

Thus, the invention provides a diagnostic method of a disorder, which involves (a) assaying the expression of a polypeptide of the present invention in cells or body fluid of an individual; (b) comparing the level of gene expression with a standard gene 15 expression level, whereby an increase or decrease in the assayed polypeptide gene expression level compared to the standard expression level is indicative of a disorder.

Moreover, polypeptides of the present invention can be used to treat disease. For example, patients can be administered a polypeptide of the present invention in an effort to replace absent or decreased levels of the polypeptide (e.g., insulin), to 20 supplement absent or decreased levels of a different polypeptide (e.g., hemoglobin S for hemoglobin B), to inhibit the activity of a polypeptide (e.g., an oncogene), to activate the activity of a polypeptide (e.g., by binding to a receptor), to reduce the activity of a membrane bound receptor by competing with it for free ligand (e.g., soluble TNF receptors used in reducing inflammation), or to bring about a desired 25 response (e.g., blood vessel growth).

Similarly, antibodies directed to a polypeptide of the present invention can also be used to treat disease. For example, administration of an antibody directed to a polypeptide of the present invention can bind and reduce overproduction of the polypeptide. Similarly, administration of an antibody can activate the polypeptide, such 30 as by binding to a polypeptide bound to a membrane (receptor).

At the very least, the polypeptides of the present invention can be used as molecular weight markers on SDS-PAGE gels or on molecular sieve gel filtration columns using methods well known to those of skill in the art. Polypeptides can also be used to raise antibodies, which in turn are used to measure protein expression from a 35 recombinant cell, as a way of assessing transformation of the host cell. Moreover, the polypeptides of the present invention can be used to test the following biological activities.

### Biological Activities

The polynucleotides and polypeptides of the present invention can be used in assays to test for one or more biological activities. If these polynucleotides and 5 polypeptides do exhibit activity in a particular assay, it is likely that these molecules may be involved in the diseases associated with the biological activity. Thus, the polynucleotides and polypeptides could be used to treat the associated disease.

### Immune Activity

A polypeptide or polynucleotide of the present invention may be useful in 10 treating deficiencies or disorders of the immune system, by activating or inhibiting the proliferation, differentiation, or mobilization (chemotaxis) of immune cells. Immune cells develop through a process called hematopoiesis, producing myeloid (platelets, red blood cells, neutrophils, and macrophages) and lymphoid (B and T lymphocytes) cells 15 from pluripotent stem cells. The etiology of these immune deficiencies or disorders may be genetic, somatic, such as cancer or some autoimmune disorders, acquired (e.g., by chemotherapy or toxins), or infectious. Moreover, a polynucleotide or polypeptide of the present invention can be used as a marker or detector of a particular immune system disease or disorder.

A polynucleotide or polypeptide of the present invention may be useful in 20 treating or detecting deficiencies or disorders of hematopoietic cells. A polypeptide or polynucleotide of the present invention could be used to increase differentiation and proliferation of hematopoietic cells, including the pluripotent stem cells, in an effort to treat those disorders associated with a decrease in certain (or many) types hematopoietic 25 cells. Examples of immunologic deficiency syndromes include, but are not limited to: blood protein disorders (e.g. agammaglobulinemia, dysgammaglobulinemia), ataxia telangiectasia, common variable immunodeficiency, Digeorge Syndrome, HIV infection, HTLV-BLV infection, leukocyte adhesion deficiency syndrome, lymphopenia, phagocyte bactericidal dysfunction, severe combined immunodeficiency 30 (SCIDs), Wiskott-Aldrich Disorder, anemia, thrombocytopenia, or hemoglobinuria.

Moreover, a polypeptide or polynucleotide of the present invention could also 35 be used to modulate hemostatic (the stopping of bleeding) or thrombolytic activity (clot formation). For example, by increasing hemostatic or thrombolytic activity, a polynucleotide or polypeptide of the present invention could be used to treat blood coagulation disorders (e.g., afibrinogenemia, factor deficiencies), blood platelet disorders (e.g. thrombocytopenia), or wounds resulting from trauma, surgery, or other causes. Alternatively, a polynucleotide or polypeptide of the present invention that can

decrease hemostatic or thrombolytic activity could be used to inhibit or dissolve clotting. These molecules could be important in the treatment of heart attacks (infarction), strokes, or scarring.

A polynucleotide or polypeptide of the present invention may also be useful in 5 treating or detecting autoimmune disorders. Many autoimmune disorders result from inappropriate recognition of self as foreign material by immune cells. This inappropriate recognition results in an immune response leading to the destruction of the host tissue. Therefore, the administration of a polypeptide or polynucleotide of the present invention that inhibits an immune response, particularly the proliferation, 10 differentiation, or chemotaxis of T-cells, may be an effective therapy in preventing autoimmune disorders.

Examples of autoimmune disorders that can be treated or detected by the present invention include, but are not limited to: Addison's Disease, hemolytic anemia, antiphospholipid syndrome, rheumatoid arthritis, dermatitis, allergic encephalomyelitis, 15 glomerulonephritis, Goodpasture's Syndrome, Graves' Disease, Multiple Sclerosis, Myasthenia Gravis, Neuritis, Ophthalmia, Bullous Pemphigoid, Pemphigus, Polyendocrinopathies, Purpura, Reiter's Disease, Stiff-Man Syndrome, Autoimmune Thyroiditis, Systemic Lupus Erythematosus, Autoimmune Pulmonary Inflammation, Guillain-Barre Syndrome, insulin dependent diabetes mellitus, and autoimmune 20 inflammatory eye disease.

Similarly, allergic reactions and conditions, such as asthma (particularly allergic asthma) or other respiratory problems, may also be treated by a polypeptide or polynucleotide of the present invention. Moreover, these molecules can be used to treat anaphylaxis, hypersensitivity to an antigenic molecule, or blood group incompatibility.

A polynucleotide or polypeptide of the present invention may also be used to 25 treat and/or prevent organ rejection or graft-versus-host disease (GVHD). Organ rejection occurs by host immune cell destruction of the transplanted tissue through an immune response. Similarly, an immune response is also involved in GVHD, but, in this case, the foreign transplanted immune cells destroy the host tissues. The 30 administration of a polypeptide or polynucleotide of the present invention that inhibits an immune response, particularly the proliferation, differentiation, or chemotaxis of T-cells, may be an effective therapy in preventing organ rejection or GVHD.

Similarly, a polypeptide or polynucleotide of the present invention may also be used to modulate inflammation. For example, the polypeptide or polynucleotide may 35 inhibit the proliferation and differentiation of cells involved in an inflammatory response. These molecules can be used to treat inflammatory conditions, both chronic and acute conditions, including inflammation associated with infection (e.g., septic

shock, sepsis, or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine induced lung injury, inflammatory bowel disease, Crohn's disease, or resulting from over production of cytokines (e.g., TNF or IL-1.)

5

### Hyperproliferative Disorders

A polypeptide or polynucleotide can be used to treat or detect hyperproliferative disorders, including neoplasms. A polypeptide or polynucleotide of the present invention may inhibit the proliferation of the disorder through direct or indirect interactions. Alternatively, a polypeptide or polynucleotide of the present invention may proliferate other cells which can inhibit the hyperproliferative disorder.

For example, by increasing an immune response, particularly increasing antigenic qualities of the hyperproliferative disorder or by proliferating, differentiating, or mobilizing T-cells, hyperproliferative disorders can be treated. This immune response may be increased by either enhancing an existing immune response, or by initiating a new immune response. Alternatively, decreasing an immune response may also be a method of treating hyperproliferative disorders, such as a chemotherapeutic agent.

Examples of hyperproliferative disorders that can be treated or detected by a polynucleotide or polypeptide of the present invention include, but are not limited to neoplasms located in the: abdomen, bone, breast, digestive system, liver, pancreas, peritoneum, endocrine glands (adrenal, parathyroid, pituitary, testicles, ovary, thymus, thyroid), eye, head and neck, nervous (central and peripheral), lymphatic system, pelvic, skin, soft tissue, spleen, thoracic, and urogenital.

Similarly, other hyperproliferative disorders can also be treated or detected by a polynucleotide or polypeptide of the present invention. Examples of such hyperproliferative disorders include, but are not limited to: hypergammaglobulinemia, lymphoproliferative disorders, paraproteinemias, purpura, sarcoidosis, Sezary Syndrome, Waldenström's Macroglobulinemia, Gaucher's Disease, histiocytosis, and any other hyperproliferative disease, besides neoplasia, located in an organ system listed above.

### Infectious Disease

A polypeptide or polynucleotide of the present invention can be used to treat or detect infectious agents. For example, by increasing the immune response, particularly increasing the proliferation and differentiation of B and/or T cells, infectious diseases

35

may be treated. The immune response may be increased by either enhancing an existing immune response, or by initiating a new immune response. Alternatively, the polypeptide or polynucleotide of the present invention may also directly inhibit the infectious agent, without necessarily eliciting an immune response.

5       Viruses are one example of an infectious agent that can cause disease or symptoms that can be treated or detected by a polynucleotide or polypeptide of the present invention. Examples of viruses, include, but are not limited to the following DNA and RNA viral families: Arbovirus, Adenoviridae, Arenaviridae, Arterivirus, Birnaviridae, Bunyaviridae, Caliciviridae, Circoviridae, Coronaviridae, Flaviviridae, 10 Hepadnaviridae (Hepatitis), Herpesviridae (such as, Cytomegalovirus, Herpes Simplex, Herpes Zoster), Mononegavirus (e.g., Paramyxoviridae, Morbillivirus, Rhabdoviridae), Orthomyxoviridae (e.g., Influenza), Papovaviridae, Parvoviridae, Picornaviridae, Poxviridae (such as Smallpox or Vaccinia), Reoviridae (e.g., Rotavirus), Retroviridae (HTLV-I, HTLV-II, Lentivirus), and Togaviridae (e.g., 15 Rubivirus). Viruses falling within these families can cause a variety of diseases or symptoms, including, but not limited to: arthritis, bronchiolitis, encephalitis, eye infections (e.g., conjunctivitis, keratitis), chronic fatigue syndrome, hepatitis (A, B, C, E, Chronic Active, Delta), meningitis, opportunistic infections (e.g., AIDS), pneumonia, Burkitt's Lymphoma, chickenpox , hemorrhagic fever, Measles, Mumps, 20 Parainfluenza, Rabies, the common cold, Polio, leukemia, Rubella, sexually transmitted diseases, skin diseases (e.g., Kaposi's, warts), and viremia. A polypeptide or polynucleotide of the present invention can be used to treat or detect any of these symptoms or diseases.

Similarly, bacterial or fungal agents that can cause disease or symptoms and that 25 can be treated or detected by a polynucleotide or polypeptide of the present invention include, but not limited to, the following Gram-Negative and Gram-positive bacterial families and fungi: Actinomycetales (e.g., Corynebacterium, Mycobacterium, Nocardia), Aspergillosis, Bacillaceae (e.g., Anthrax, Clostridium), Bacteroidaceae, Blastomycosis, Bordetella, Borrelia, Brucellosis, Candidiasis, Campylobacter, 30 Coccidioidomycosis, Cryptococcosis, Dermatocycoses, Enterobacteriaceae (Klebsiella, Salmonella, Serratia, Yersinia), Erysipelothrix, Helicobacter, Legionellosis, Leptospirosis, Listeria, Mycoplasmatales, Neisseriaceae (e.g., Acinetobacter, Gonorrhea, Menigococcal), Pasteurellacea Infections (e.g., Actinobacillus, Hearnophilus, Pasteurella), Pseudomonas, Rickettsiaceae, Chlamydiaceae, Syphilis, 35 and Staphylococcal. These bacterial or fungal families can cause the following diseases or symptoms, including, but not limited to: bacteremia, endocarditis, eye infections (conjunctivitis, tuberculosis, uveitis), gingivitis, opportunistic infections (e.g., AIDS

related infections), paronychia, prostheses-related infections, Reiter's Disease, respiratory tract infections, such as Whooping Cough or Empyema, sepsis, Lyme Disease, Cat-Scratch Disease, Dysentery, Paratyphoid Fever, food poisoning, Typhoid, pneumonia, Gonorrhea, meningitis, Chlamydia, Syphilis, Diphtheria, 5 Leprosy, Paratuberculosis, Tuberculosis, Lupus, Botulism, gangrene, tetanus, impetigo, Rheumatic Fever, Scarlet Fever, sexually transmitted diseases, skin diseases (e.g., cellulitis, dermatocycoses), toxemia, urinary tract infections, wound infections. A polypeptide or polynucleotide of the present invention can be used to treat or detect any of these symptoms or diseases.

10 Moreover, parasitic agents causing disease or symptoms that can be treated or detected by a polynucleotide or polypeptide of the present invention include, but not limited to, the following families: Amebiasis, Babesiosis, Coccidioidosis, Cryptosporidiosis, Dientamoebiasis, Dourine, Ectoparasitic, Giardiasis, Helminthiasis, Leishmaniasis, Theileriasis, Toxoplasmosis, Trypanosomiasis, and Trichomonas. 15 These parasites can cause a variety of diseases or symptoms, including, but not limited to: Scabies, Trombiculiasis, eye infections, intestinal disease (e.g., dysentery, giardiasis), liver disease, lung disease, opportunistic infections (e.g., AIDS related), Malaria, pregnancy complications, and toxoplasmosis. A polypeptide or polynucleotide of the present invention can be used to treat or detect any of these symptoms or 20 diseases.

25 Preferably, treatment using a polypeptide or polynucleotide of the present invention could either be by administering an effective amount of a polypeptide to the patient, or by removing cells from the patient, supplying the cells with a polynucleotide of the present invention, and returning the engineered cells to the patient (ex vivo therapy). Moreover, the polypeptide or polynucleotide of the present invention can be used as an antigen in a vaccine to raise an immune response against infectious disease.

### Regeneration

30 A polynucleotide or polypeptide of the present invention can be used to differentiate, proliferate, and attract cells, leading to the regeneration of tissues. (See, Science 276:59-87 (1997).) The regeneration of tissues could be used to repair, replace, or protect tissue damaged by congenital defects, trauma (wounds, burns, incisions, or ulcers), age, disease (e.g. osteoporosis, osteoarthritis, periodontal disease, liver failure), surgery, including cosmetic plastic surgery, fibrosis, reperfusion 35 injury, or systemic cytokine damage.

Tissues that could be regenerated using the present invention include organs (e.g., pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal

or cardiac), vascular (including vascular endothelium), nervous, hematopoietic, and skeletal (bone, cartilage, tendon, and ligament) tissue. Preferably, regeneration occurs without or decreased scarring. Regeneration also may include angiogenesis.

Moreover, a polynucleotide or polypeptide of the present invention may increase regeneration of tissues difficult to heal. For example, increased tendon/ligament regeneration would quicken recovery time after damage. A polynucleotide or polypeptide of the present invention could also be used prophylactically in an effort to avoid damage. Specific diseases that could be treated include of tendinitis, carpal tunnel syndrome, and other tendon or ligament defects. A further example of tissue regeneration of non-healing wounds includes pressure ulcers, ulcers associated with vascular insufficiency, surgical, and traumatic wounds.

Similarly, nerve and brain tissue could also be regenerated by using a polynucleotide or polypeptide of the present invention to proliferate and differentiate nerve cells. Diseases that could be treated using this method include central and peripheral nervous system diseases, neuropathies, or mechanical and traumatic disorders (e.g., spinal cord disorders, head trauma, cerebrovascular disease, and stroke). Specifically, diseases associated with peripheral nerve injuries, peripheral neuropathy (e.g., resulting from chemotherapy or other medical therapies), localized neuropathies, and central nervous system diseases (e.g., Alzheimer's disease, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome), could all be treated using the polynucleotide or polypeptide of the present invention.

### **Chemotaxis**

A polynucleotide or polypeptide of the present invention may have chemotaxis activity. A chemotactic molecule attracts or mobilizes cells (e.g., monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells) to a particular site in the body, such as inflammation, infection, or site of hyperproliferation. The mobilized cells can then fight off and/or heal the particular trauma or abnormality.

A polynucleotide or polypeptide of the present invention may increase chemotactic activity of particular cells. These chemotactic molecules can then be used to treat inflammation, infection, hyperproliferative disorders, or any immune system disorder by increasing the number of cells targeted to a particular location in the body. For example, chemotactic molecules can be used to treat wounds and other trauma to tissues by attracting immune cells to the injured location. Chemotactic molecules of the present invention can also attract fibroblasts, which can be used to treat wounds.

It is also contemplated that a polynucleotide or polypeptide of the present invention may inhibit chemotactic activity. These molecules could also be used to treat disorders. Thus, a polynucleotide or polypeptide of the present invention could be used as an inhibitor of chemotaxis.

5

### Binding Activity

A polypeptide of the present invention may be used to screen for molecules that bind to the polypeptide or for molecules to which the polypeptide binds. The binding of the polypeptide and the molecule may activate (agonist), increase, inhibit 10 (antagonist), or decrease activity of the polypeptide or the molecule bound. Examples of such molecules include antibodies, oligonucleotides, proteins (e.g., receptors), or small molecules.

Preferably, the molecule is closely related to the natural ligand of the 15 polypeptide, e.g., a fragment of the ligand, or a natural substrate, a ligand, a structural or functional mimetic. (See, Coligan et al., Current Protocols in Immunology 1(2):Chapter 5 (1991).) Similarly, the molecule can be closely related to the natural receptor to which the polypeptide binds, or at least, a fragment of the receptor capable of being bound by the polypeptide (e.g., active site). In either case, the molecule can be rationally designed using known techniques.

20 Preferably, the screening for these molecules involves producing appropriate cells which express the polypeptide, either as a secreted protein or on the cell membrane. Preferred cells include cells from mammals, yeast, Drosophila, or *E. coli*. Cells expressing the polypeptide (or cell membrane containing the expressed 25 polypeptide) are then preferably contacted with a test compound potentially containing the molecule to observe binding, stimulation, or inhibition of activity of either the polypeptide or the molecule.

The assay may simply test binding of a candidate compound to the polypeptide, wherein binding is detected by a label, or in an assay involving competition with a labeled competitor. Further, the assay may test whether the candidate compound results 30 in a signal generated by binding to the polypeptide.

Alternatively, the assay can be carried out using cell-free preparations, polypeptide/molecule affixed to a solid support, chemical libraries, or natural product mixtures. The assay may also simply comprise the steps of mixing a candidate compound with a solution containing a polypeptide, measuring polypeptide/molecule 35 activity or binding, and comparing the polypeptide/molecule activity or binding to a standard.

Preferably, an ELISA assay can measure polypeptide level or activity in a sample (e.g., biological sample) using a monoclonal or polyclonal antibody. The antibody can measure polypeptide level or activity by either binding, directly or indirectly, to the polypeptide or by competing with the polypeptide for a substrate.

5 All of these above assays can be used as diagnostic or prognostic markers. The molecules discovered using these assays can be used to treat disease or to bring about a particular result in a patient (e.g., blood vessel growth) by activating or inhibiting the polypeptide/molecule. Moreover, the assays can discover agents which may inhibit or enhance the production of the polypeptide from suitably manipulated cells or tissues.

10 Therefore, the invention includes a method of identifying compounds which bind to a polypeptide of the invention comprising the steps of: (a) incubating a candidate binding compound with a polypeptide of the invention; and (b) determining if binding has occurred. Moreover, the invention includes a method of identifying agonists/antagonists comprising the steps of: (a) incubating a candidate compound with 15 a polypeptide of the invention, (b) assaying a biological activity , and (b) determining if a biological activity of the polypeptide has been altered.

### Other Activities

A polypeptide or polynucleotide of the present invention may also increase or 20 decrease the differentiation or proliferation of embryonic stem cells, besides, as discussed above, hematopoietic lineage.

A polypeptide or polynucleotide of the present invention may also be used to modulate mammalian characteristics, such as body height, weight, hair color, eye color, skin, percentage of adipose tissue, pigmentation, size, and shape (e.g., cosmetic 25 surgery). Similarly, a polypeptide or polynucleotide of the present invention may be used to modulate mammalian metabolism affecting catabolism, anabolism, processing, utilization, and storage of energy.

A polypeptide or polynucleotide of the present invention may be used to change a mammal's mental state or physical state by influencing biorhythms, circadian 30 rhythms, depression (including depressive disorders), tendency for violence, tolerance for pain, reproductive capabilities (preferably by Activin or Inhibin-like activity), hormonal or endocrine levels, appetite, libido, memory, stress, or other cognitive qualities.

A polypeptide or polynucleotide of the present invention may also be used as a 35 food additive or preservative, such as to increase or decrease storage capabilities, fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors or other nutritional components.

**Other Preferred Embodiments**

Other preferred embodiments of the claimed invention include an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 50 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1.

Also preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of positions beginning with the nucleotide at about the position of the 5' Nucleotide of the Clone Sequence and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

Also preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of positions beginning with the nucleotide at about the position of the 5' Nucleotide of the Start Codon and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

Similarly preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of positions beginning with the nucleotide at about the position of the 5' Nucleotide of the First Amino Acid of the Signal Peptide and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 150 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X.

Further preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 500 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X.

A further preferred embodiment is a nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the nucleotide sequence of SEQ ID NO:X beginning with the nucleotide at about the position of the 5' Nucleotide of the First Amino Acid of the Signal Peptide and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the complete nucleotide sequence of SEQ ID NO:X.

Also preferred is an isolated nucleic acid molecule which hybridizes under 5 stringent hybridization conditions to a nucleic acid molecule, wherein said nucleic acid molecule which hybridizes does not hybridize under stringent hybridization conditions to a nucleic acid molecule having a nucleotide sequence consisting of only A residues or of only T residues.

Also preferred is a composition of matter comprising a DNA molecule which 10 comprises a human cDNA clone identified by a cDNA Clone Identifier in Table 1, which DNA molecule is contained in the material deposited with the American Type Culture Collection and given the ATCC Deposit Number shown in Table 1 for said cDNA Clone Identifier.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide 15 sequence which is at least 95% identical to a sequence of at least 50 contiguous nucleotides in the nucleotide sequence of a human cDNA clone identified by a cDNA Clone Identifier in Table 1, which DNA molecule is contained in the deposit given the ATCC Deposit Number shown in Table 1.

Also preferred is an isolated nucleic acid molecule, wherein said sequence of at 20 least 50 contiguous nucleotides is included in the nucleotide sequence of the complete open reading frame sequence encoded by said human cDNA clone.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to sequence of at least 150 contiguous nucleotides in the nucleotide sequence encoded by said human cDNA clone.

25 A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to sequence of at least 500 contiguous nucleotides in the nucleotide sequence encoded by said human cDNA clone.

A further preferred embodiment is an isolated nucleic acid molecule comprising 30 a nucleotide sequence which is at least 95% identical to the complete nucleotide sequence encoded by said human cDNA clone.

A further preferred embodiment is a method for detecting in a biological sample a nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer 35 as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1; which method.

comprises a step of comparing a nucleotide sequence of at least one nucleic acid molecule in said sample with a sequence selected from said group and determining whether the sequence of said nucleic acid molecule in said sample is at least 95% identical to said selected sequence.

5 Also preferred is the above method wherein said step of comparing sequences comprises determining the extent of nucleic acid hybridization between nucleic acid molecules in said sample and a nucleic acid molecule comprising said sequence selected from said group. Similarly, also preferred is the above method wherein said step of comparing sequences is performed by comparing the nucleotide sequence determined  
10 from a nucleic acid molecule in said sample with said sequence selected from said group. The nucleic acid molecules can comprise DNA molecules or RNA molecules.

A further preferred embodiment is a method for identifying the species, tissue or cell type of a biological sample which method comprises a step of detecting nucleic acid molecules in said sample, if any, comprising a nucleotide sequence that is at least 95%  
15 identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

20 The method for identifying the species, tissue or cell type of a biological sample can comprise a step of detecting nucleic acid molecules comprising a nucleotide sequence in a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from said group.

25 Also preferred is a method for diagnosing in a subject a pathological condition associated with abnormal structure or expression of a gene encoding a secreted protein identified in Table 1, which method comprises a step of detecting in a biological sample obtained from said subject nucleic acid molecules, if any, comprising a nucleotide sequence that is at least 95% identical to a sequence of at least 50 contiguous  
30 nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

35 The method for diagnosing a pathological condition can comprise a step of detecting nucleic acid molecules comprising a nucleotide sequence in a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95%

identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from said group.

Also preferred is a composition of matter comprising isolated nucleic acid molecules wherein the nucleotide sequences of said nucleic acid molecules comprise a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1. The nucleic acid molecules can comprise DNA molecules or RNA molecules.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 90% identical to a sequence of at least about 10 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1.

Also preferred is a polypeptide, wherein said sequence of contiguous amino acids is included in the amino acid sequence of SEQ ID NO:Y in the range of positions beginning with the residue at about the position of the First Amino Acid of the Secreted Portion and ending with the residue at about the Last Amino Acid of the Open Reading Frame as set forth for SEQ ID NO:Y in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 30 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 100 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to the complete amino acid sequence of SEQ ID NO:Y.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 90% identical to a sequence of at least about 10 contiguous amino acids in the complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is a polypeptide wherein said sequence of contiguous amino acids is included in the amino acid sequence of a secreted portion of the secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 30 contiguous amino acids in the amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 100 contiguous amino acids in the amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to the amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Further preferred is an isolated antibody which binds specifically to a polypeptide comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Further preferred is a method for detecting in a biological sample a polypeptide comprising an amino acid sequence which is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1; which method comprises a step of comparing an amino acid sequence of at least one polypeptide molecule in said sample with a sequence selected from said group and determining whether the sequence of said polypeptide molecule in said sample is at least 90% identical to said sequence of at least 10 contiguous amino acids.

Also preferred is the above method wherein said step of comparing an amino acid sequence of at least one polypeptide molecule in said sample with a sequence selected from said group comprises determining the extent of specific binding of polypeptides in said sample to an antibody which binds specifically to a polypeptide

comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is the above method wherein said step of comparing sequences is performed by comparing the amino acid sequence determined from a polypeptide molecule in said sample with said sequence selected from said group.

Also preferred is a method for identifying the species, tissue or cell type of a biological sample which method comprises a step of detecting polypeptide molecules in said sample, if any, comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is the above method for identifying the species, tissue or cell type of a biological sample, which method comprises a step of detecting polypeptide molecules comprising an amino acid sequence in a panel of at least two amino acid sequences, wherein at least one sequence in said panel is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the above group.

Also preferred is a method for diagnosing in a subject a pathological condition associated with abnormal structure or expression of a gene encoding a secreted protein identified in Table 1, which method comprises a step of detecting in a biological sample obtained from said subject polypeptide molecules comprising an amino acid sequence in a panel of at least two amino acid sequences, wherein at least one sequence in said panel is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

In any of these methods, the step of detecting said polypeptide molecules includes using an antibody.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a nucleotide sequence encoding a polypeptide wherein said polypeptide comprises an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated nucleic acid molecule, wherein said nucleotide sequence encoding a polypeptide has been optimized for expression of said polypeptide in a prokaryotic host.

Also preferred is an isolated nucleic acid molecule, wherein said polypeptide comprises an amino acid sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Further preferred is a method of making a recombinant vector comprising inserting any of the above isolated nucleic acid molecule into a vector. Also preferred is the recombinant vector produced by this method. Also preferred is a method of making a recombinant host cell comprising introducing the vector into a host cell, as well as the recombinant host cell produced by this method.

Also preferred is a method of making an isolated polypeptide comprising culturing this recombinant host cell under conditions such that said polypeptide is expressed and recovering said polypeptide. Also preferred is this method of making an isolated polypeptide, wherein said recombinant host cell is a eukaryotic cell and said polypeptide is a secreted portion of a human secreted protein comprising an amino acid sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y beginning with the residue at the position of the First Amino Acid of the Secreted Portion of SEQ ID NO:Y wherein Y is an integer set forth in Table 1 and said position of the First Amino Acid of the Secreted Portion of SEQ ID NO:Y is defined in Table 1; and an amino acid sequence of a secreted portion of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1. The isolated polypeptide produced by this method is also preferred.

Also preferred is a method of treatment of an individual in need of an increased level of a secreted protein activity, which method comprises administering to such an individual a pharmaceutical composition comprising an amount of an isolated polypeptide, polynucleotide, or antibody of the claimed invention effective to increase 5 the level of said protein activity in said individual.

Having generally described the invention, the same will be more readily understood by reference to the following examples, which are provided by way of illustration and are not intended as limiting.

10

### Examples

#### Example 1: Isolation of a Selected cDNA Clone From the Deposited Sample

Each cDNA clone in a cited ATCC deposit is contained in a plasmid vector. 15 Table 1 identifies the vectors used to construct the cDNA library from which each clone was isolated. In many cases, the vector used to construct the library is a phage vector from which a plasmid has been excised. The table immediately below correlates the related plasmid for each phage vector used in constructing the cDNA library. For example, where a particular clone is identified in Table 1 as being isolated in the vector 20 "Lambda Zap," the corresponding deposited clone is in "pBluescript."

<u>Vector Used to Construct Library</u>	<u>Corresponding Deposited Plasmid</u>
Lambda Zap	pBluescript (pBS)
Uni-Zap XR	pBluescript (pBS)
Zap Express	pBK
25 lafmid BA	plafmid BA
pSport1	pSport1
pCMVSport 2.0	pCMVSport 2.0
pCMVSport 3.0	pCMVSport 3.0
pCR®2.1	pCR®2.1

30 Vectors Lambda Zap (U.S. Patent Nos. 5,128,256 and 5,286,636), Uni-Zap  
XR (U.S. Patent Nos. 5,128, 256 and 5,286,636), Zap Express (U.S. Patent Nos.  
5,128,256 and 5,286,636), pBluescript (pBS) (Short, J. M. et al., Nucleic Acids Res.  
16:7583-7600 (1988); Alting-Mees, M. A. and Short, J. M., Nucleic Acids Res.  
17:9494 (1989)) and pBK (Alting-Mees, M. A. et al., Strategies 5:58-61 (1992)) are  
35 commercially available from Stratagene Cloning Systems, Inc., 11011 N. Torrey Pines  
Road, La Jolla, CA, 92037. pBS contains an ampicillin resistance gene and pBK  
contains a neomycin resistance gene. Both can be transformed into E. coli strain XL-1

Blue, also available from Stratagene. pBS comes in 4 forms SK+, SK-, KS+ and KS. The S and K refers to the orientation of the polylinker to the T7 and T3 primer sequences which flank the polylinker region ("S" is for SacI and "K" is for KpnI which are the first sites on each respective end of the linker). "+" or "-" refer to the orientation of the f1 origin of replication ("ori"), such that in one orientation, single stranded rescue initiated from the f1 ori generates sense strand DNA and in the other, antisense.

Vectors pSport1, pCMVSport 2.0 and pCMVSport 3.0, were obtained from Life Technologies, Inc., P. O. Box 6009, Gaithersburg, MD 20897. All Sport vectors contain an ampicillin resistance gene and may be transformed into E. coli strain 10 DH10B, also available from Life Technologies. (See, for instance, Gruber, C. E., et al., Focus 15:59 (1993).) Vector lafmid BA (Bento Soares, Columbia University, NY) contains an ampicillin resistance gene and can be transformed into E. coli strain XL-1 Blue. Vector pCR®2.1, which is available from Invitrogen, 1600 Faraday Avenue, Carlsbad, CA 92008, contains an ampicillin resistance gene and may be transformed 15 into E. coli strain DH10B, available from Life Technologies. (See, for instance, Clark, J. M., Nuc. Acids Res. 16:9677-9686 (1988) and Mead, D. et al., Bio/Technology 9: (1991).) Preferably, a polynucleotide of the present invention does not comprise the phage vector sequences identified for the particular clone in Table 1, as well as the corresponding plasmid vector sequences designated above.

20 The deposited material in the sample assigned the ATCC Deposit Number cited in Table 1 for any given cDNA clone also may contain one or more additional plasmids, each comprising a cDNA clone different from that given clone. Thus, deposits sharing the same ATCC Deposit Number contain at least a plasmid for each cDNA clone identified in Table 1. Typically, each ATCC deposit sample cited in Table 1 comprises 25 a mixture of approximately equal amounts (by weight) of about 50 plasmid DNAs, each containing a different cDNA clone; but such a deposit sample may include plasmids for more or less than 50 cDNA clones, up to about 500 cDNA clones.

Two approaches can be used to isolate a particular clone from the deposited 30 sample of plasmid DNAs cited for that clone in Table 1. First, a plasmid is directly isolated by screening the clones using a polynucleotide probe corresponding to SEQ ID NO:X.

Particularly, a specific polynucleotide with 30-40 nucleotides is synthesized using an Applied Biosystems DNA synthesizer according to the sequence reported. The oligonucleotide is labeled, for instance, with  $^{32}\text{P}$ - $\gamma$ -ATP using T4 polynucleotide 35 kinase and purified according to routine methods. (E.g., Maniatis et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Press, Cold Spring, NY (1982).)

The plasmid mixture is transformed into a suitable host, as indicated above (such as XL-1 Blue (Stratagene)) using techniques known to those of skill in the art, such as those provided by the vector supplier or in related publications or patents cited above. The transformants are plated on 1.5% agar plates (containing the appropriate selection agent, e.g., ampicillin) to a density of about 150 transformants (colonies) per plate. These plates are screened using Nylon membranes according to routine methods for bacterial colony screening (e.g., Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd Edit., (1989), Cold Spring Harbor Laboratory Press, pages 1.93 to 1.104), or other techniques known to those of skill in the art.

Alternatively, two primers of 17-20 nucleotides derived from both ends of the SEQ ID NO:X (i.e., within the region of SEQ ID NO:X bounded by the 5' NT and the 3' NT of the clone defined in Table 1) are synthesized and used to amplify the desired cDNA using the deposited cDNA plasmid as a template. The polymerase chain reaction is carried out under routine conditions, for instance, in 25 µl of reaction mixture with 0.5 ug of the above cDNA template. A convenient reaction mixture is 1.5-5 mM MgCl<sub>2</sub>, 0.01% (w/v) gelatin, 20 µM each of dATP, dCTP, dGTP, dTTP, 25 pmol of each primer and 0.25 Unit of Taq polymerase. Thirty five cycles of PCR (denaturation at 94°C for 1 min; annealing at 55°C for 1 min; elongation at 72°C for 1 min) are performed with a Perkin-Elmer Cetus automated thermal cycler. The amplified product is analyzed by agarose gel electrophoresis and the DNA band with expected molecular weight is excised and purified. The PCR product is verified to be the selected sequence by subcloning and sequencing the DNA product.

Several methods are available for the identification of the 5' or 3' non-coding portions of a gene which may not be present in the deposited clone. These methods include but are not limited to, filter probing, clone enrichment using specific probes, and protocols similar or identical to 5' and 3' "RACE" protocols which are well known in the art. For instance, a method similar to 5' RACE is available for generating the missing 5' end of a desired full-length transcript. (Fromont-Racine et al., Nucleic Acids Res. 21(7):1683-1684 (1993).)

Briefly, a specific RNA oligonucleotide is ligated to the 5' ends of a population of RNA presumably containing full-length gene RNA transcripts. A primer set containing a primer specific to the ligated RNA oligonucleotide and a primer specific to a known sequence of the gene of interest is used to PCR amplify the 5' portion of the desired full-length gene. This amplified product may then be sequenced and used to generate the full length gene.

This above method starts with total RNA isolated from the desired source, although poly-A+ RNA can be used. The RNA preparation can then be treated with phosphatase if necessary to eliminate 5' phosphate groups on degraded or damaged RNA which may interfere with the later RNA ligase step. The phosphatase should then 5 be inactivated and the RNA treated with tobacco acid pyrophosphatase in order to remove the cap structure present at the 5' ends of messenger RNAs. This reaction leaves a 5' phosphate group at the 5' end of the cap cleaved RNA which can then be ligated to an RNA oligonucleotide using T4 RNA ligase.

This modified RNA preparation is used as a template for first strand cDNA 10 synthesis using a gene specific oligonucleotide. The first strand synthesis reaction is used as a template for PCR amplification of the desired 5' end using a primer specific to the ligated RNA oligonucleotide and a primer specific to the known sequence of the gene of interest. The resultant product is then sequenced and analyzed to confirm that the 5' end sequence belongs to the desired gene.

15

**Example 2: Isolation of Genomic Clones Corresponding to a Polynucleotide**

A human genomic P1 library (Genomic Systems, Inc.) is screened by PCR 20 using primers selected for the cDNA sequence corresponding to SEQ ID NO:X., according to the method described in Example 1. (See also, Sambrook.)

**Example 3: Tissue Distribution of Polypeptide**

Tissue distribution of mRNA expression of polynucleotides of the present invention is determined using protocols for Northern blot analysis, described by, 25 among others, Sambrook et al. For example, a cDNA probe produced by the method described in Example 1 is labeled with P<sup>32</sup> using the rediprime™ DNA labeling system (Amersham Life Science), according to manufacturer's instructions. After labeling, the probe is purified using CHROMA SPIN-100™ column (Clontech Laboratories, Inc.), according to manufacturer's protocol number PT1200-1. The purified labeled probe is 30 then used to examine various human tissues for mRNA expression.

Multiple Tissue Northern (MTN) blots containing various human tissues (H) or human immune system tissues (IM) (Clontech) are examined with the labeled probe using ExpressHyb™ hybridization solution (Clontech) according to manufacturer's protocol number PT1190-1. Following hybridization and washing, the blots are 35 mounted and exposed to film at -70°C overnight, and the films developed according to standard procedures.

**Example 4: Chromosomal Mapping of the Polynucleotides**

An oligonucleotide primer set is designed according to the sequence at the 5' end of SEQ ID NO:X. This primer preferably spans about 100 nucleotides. This 5 primer set is then used in a polymerase chain reaction under the following set of conditions : 30 seconds, 95°C; 1 minute, 56°C; 1 minute, 70°C. This cycle is repeated 32 times followed by one 5 minute cycle at 70°C. Human, mouse, and hamster DNA is used as template in addition to a somatic cell hybrid panel containing individual chromosomes or chromosome fragments (Bios, Inc). The reactions is analyzed on 10 either 8% polyacrylamide gels or 3.5 % agarose gels. Chromosome mapping is determined by the presence of an approximately 100 bp PCR fragment in the particular somatic cell hybrid.

**Example 5: Bacterial Expression of a Polypeptide**

15 A polynucleotide encoding a polypeptide of the present invention is amplified using PCR oligonucleotide primers corresponding to the 5' and 3' ends of the DNA sequence, as outlined in Example 1, to synthesize insertion fragments. The primers used to amplify the cDNA insert should preferably contain restriction sites, such as BamHI and XbaI, at the 5' end of the primers in order to clone the amplified product 20 into the expression vector. For example, BamHI and XbaI correspond to the restriction enzyme sites on the bacterial expression vector pQE-9. (Qiagen, Inc., Chatsworth, CA). This plasmid vector encodes antibiotic resistance ( $\text{Amp}^r$ ), a bacterial origin of replication (ori), an IPTG-regulatable promoter/operator (P/O), a ribosome binding site (RBS), a 6-histidine tag (6-His), and restriction enzyme cloning sites.

25 The pQE-9 vector is digested with BamHI and XbaI and the amplified fragment is ligated into the pQE-9 vector maintaining the reading frame initiated at the bacterial RBS. The ligation mixture is then used to transform the *E. coli* strain M15/rep4 (Qiagen, Inc.) which contains multiple copies of the plasmid pREP4, which expresses the lacI repressor and also confers kanamycin resistance ( $\text{Kan}^r$ ). Transformants are 30 identified by their ability to grow on LB plates and ampicillin/kanamycin resistant colonies are selected. Plasmid DNA is isolated and confirmed by restriction analysis.

Clones containing the desired constructs are grown overnight (O/N) in liquid culture in LB media supplemented with both Amp (100 ug/ml) and Kan (25 ug/ml). The O/N culture is used to inoculate a large culture at a ratio of 1:100 to 1:250. The 35 cells are grown to an optical density 600 (O.D.  $^{600}$ ) of between 0.4 and 0.6. IPTG

(Isopropyl-B-D-thiogalacto pyranoside) is then added to a final concentration of 1 mM. IPTG induces by inactivating the lacI repressor, clearing the P/O leading to increased gene expression.

Cells are grown for an extra 3 to 4 hours. Cells are then harvested by centrifugation (20 mins at 6000Xg). The cell pellet is solubilized in the chaotropic agent 6 Molar Guanidine HCl by stirring for 3-4 hours at 4°C. The cell debris is removed by centrifugation, and the supernatant containing the polypeptide is loaded onto a nickel-nitrilo-tri-acetic acid ("Ni-NTA") affinity resin column (available from QIAGEN, Inc., *supra*). Proteins with a 6 x His tag bind to the Ni-NTA resin with high affinity and can be purified in a simple one-step procedure (for details see: The QIAexpressionist (1995) QIAGEN, Inc., *supra*).

Briefly, the supernatant is loaded onto the column in 6 M guanidine-HCl, pH 8, the column is first washed with 10 volumes of 6 M guanidine-HCl, pH 8, then washed with 10 volumes of 6 M guanidine-HCl pH 6, and finally the polypeptide is eluted with 6 M guanidine-HCl, pH 5.

The purified protein is then renatured by dialyzing it against phosphate-buffered saline (PBS) or 50 mM Na-acetate, pH 6 buffer plus 200 mM NaCl. Alternatively, the protein can be successfully refolded while immobilized on the Ni-NTA column. The recommended conditions are as follows: renature using a linear 6M-1M urea gradient in 500 mM NaCl, 20% glycerol, 20 mM Tris/HCl pH 7.4, containing protease inhibitors. The renaturation should be performed over a period of 1.5 hours or more. After renaturation the proteins are eluted by the addition of 250 mM immidazole. Immidazole is removed by a final dialyzing step against PBS or 50 mM sodium acetate pH 6 buffer plus 200 mM NaCl. The purified protein is stored at 4°C or frozen at -80°C.

In addition to the above expression vector, the present invention further includes an expression vector comprising phage operator and promoter elements operatively linked to a polynucleotide of the present invention, called pHE4a. (ATCC Accession Number 209645, deposited on February 25, 1998.) This vector contains: 1) a neomycinphosphotransferase gene as a selection marker, 2) an E. coli origin of replication, 3) a T5 phage promoter sequence, 4) two lac operator sequences, 5) a Shine-Delgarno sequence, and 6) the lactose operon repressor gene (lacIq). The origin of replication (oriC) is derived from pUC19 (LTI, Gaithersburg, MD). The promoter sequence and operator sequences are made synthetically.

DNA can be inserted into the pHEa by restricting the vector with NdeI and XbaI, BamHI, XhoI, or Asp718, running the restricted product on a gel, and isolating the larger fragment (the stuffer fragment should be about 310 base pairs). The DNA

insert is generated according to the PCR protocol described in Example 1, using PCR primers having restriction sites for NdeI (5' primer) and XbaI, BamHI, XhoI, or Asp718 (3' primer). The PCR insert is gel purified and restricted with compatible enzymes. The insert and vector are ligated according to standard protocols.

5        The engineered vector could easily be substituted in the above protocol to express protein in a bacterial system.

**Example 6: Purification of a Polypeptide from an Inclusion Body**

10      The following alternative method can be used to purify a polypeptide expressed in *E coli* when it is present in the form of inclusion bodies. Unless otherwise specified, all of the following steps are conducted at 4-10°C.

15      Upon completion of the production phase of the *E. coli* fermentation, the cell culture is cooled to 4-10°C and the cells harvested by continuous centrifugation at 15,000 rpm (Heraeus Sepatech). On the basis of the expected yield of protein per unit weight of cell paste and the amount of purified protein required, an appropriate amount of cell paste, by weight, is suspended in a buffer solution containing 100 mM Tris, 50 mM EDTA, pH 7.4. The cells are dispersed to a homogeneous suspension using a high shear mixer.

20      The cells are then lysed by passing the solution through a microfluidizer (Microfluidics, Corp. or APV Gaulin, Inc.) twice at 4000-6000 psi. The homogenate is then mixed with NaCl solution to a final concentration of 0.5 M NaCl, followed by centrifugation at 7000 xg for 15 min. The resultant pellet is washed again using 0.5M NaCl, 100 mM Tris, 50 mM EDTA, pH 7.4.

25      The resulting washed inclusion bodies are solubilized with 1.5 M guanidine hydrochloride (GuHCl) for 2-4 hours. After 7000 xg centrifugation for 15 min., the pellet is discarded and the polypeptide containing supernatant is incubated at 4°C overnight to allow further GuHCl extraction.

30      Following high speed centrifugation (30,000 xg) to remove insoluble particles, the GuHCl solubilized protein is refolded by quickly mixing the GuHCl extract with 20 volumes of buffer containing 50 mM sodium, pH 4.5, 150 mM NaCl, 2 mM EDTA by vigorous stirring. The refolded diluted protein solution is kept at 4°C without mixing for 12 hours prior to further purification steps.

To clarify the refolded polypeptide solution, a previously prepared tangential filtration unit equipped with 0.16 µm membrane filter with appropriate surface area

(e.g., Filtron), equilibrated with 40 mM sodium acetate, pH 6.0 is employed. The filtered sample is loaded onto a cation exchange resin (e.g., Poros HS-50, Perseptive Biosystems). The column is washed with 40 mM sodium acetate, pH 6.0 and eluted with 250 mM, 500 mM, 1000 mM, and 1500 mM NaCl in the same buffer, in a 5 stepwise manner. The absorbance at 280 nm of the effluent is continuously monitored. Fractions are collected and further analyzed by SDS-PAGE.

Fractions containing the polypeptide are then pooled and mixed with 4 volumes of water. The diluted sample is then loaded onto a previously prepared set of tandem columns of strong anion (Poros HQ-50, Perseptive Biosystems) and weak anion 10 (Poros CM-20, Perseptive Biosystems) exchange resins. The columns are equilibrated with 40 mM sodium acetate, pH 6.0. Both columns are washed with 40 mM sodium acetate, pH 6.0, 200 mM NaCl. The CM-20 column is then eluted using a 10 column volume linear gradient ranging from 0.2 M NaCl, 50 mM sodium acetate, pH 6.0 to 1.0 M NaCl, 50 mM sodium acetate, pH 6.5. Fractions are collected under constant A<sub>280</sub> 15 monitoring of the effluent. Fractions containing the polypeptide (determined, for instance, by 16% SDS-PAGE) are then pooled.

The resultant polypeptide should exhibit greater than 95% purity after the above refolding and purification steps. No major contaminant bands should be observed from Commassie blue stained 16% SDS-PAGE gel when 5 µg of purified protein is loaded. 20 The purified protein can also be tested for endotoxin/LPS contamination, and typically the LPS content is less than 0.1 ng/ml according to LAL assays.

#### Example 7: Cloning and Expression of a Polypeptide in a Baculovirus

##### Expression System

In this example, the plasmid shuttle vector pA2 is used to insert a polynucleotide 25 into a baculovirus to express a polypeptide. This expression vector contains the strong polyhedrin promoter of the *Autographa californica* nuclear polyhedrosis virus (AcMNPV) followed by convenient restriction sites such as BamHI, Xba I and Asp718. The polyadenylation site of the simian virus 40 ("SV40") is used for efficient 30 polyadenylation. For easy selection of recombinant virus, the plasmid contains the beta-galactosidase gene from *E. coli* under control of a weak Drosophila promoter in the same orientation, followed by the polyadenylation signal of the polyhedrin gene. The inserted genes are flanked on both sides by viral sequences for cell-mediated homologous recombination with wild-type viral DNA to generate a viable virus that 35 express the cloned polynucleotide.

Many other baculovirus vectors can be used in place of the vector above, such as pAc373, pVL941, and pAcIM1, as one skilled in the art would readily appreciate, as long as the construct provides appropriately located signals for transcription, translation, secretion and the like, including a signal peptide and an in-frame AUG as required. Such vectors are described, for instance, in Luckow et al., *Virology* 170:31-39 (1989).

Specifically, the cDNA sequence contained in the deposited clone, including the AUG initiation codon and the naturally associated leader sequence identified in Table 1, is amplified using the PCR protocol described in Example 1. If the naturally occurring signal sequence is used to produce the secreted protein, the pA2 vector does not need a second signal peptide. Alternatively, the vector can be modified (pA2 GP) to include a baculovirus leader sequence, using the standard methods described in Summers et al., "A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures," Texas Agricultural Experimental Station Bulletin No. 1555 (1987).

The amplified fragment is isolated from a 1% agarose gel using a commercially available kit ("Geneclean," BIO 101 Inc., La Jolla, Ca.). The fragment then is digested with appropriate restriction enzymes and again purified on a 1% agarose gel.

The plasmid is digested with the corresponding restriction enzymes and optionally, can be dephosphorylated using calf intestinal phosphatase, using routine procedures known in the art. The DNA is then isolated from a 1% agarose gel using a commercially available kit ("Geneclean" BIO 101 Inc., La Jolla, Ca.).

The fragment and the dephosphorylated plasmid are ligated together with T4 DNA ligase. *E. coli* HB101 or other suitable *E. coli* hosts such as XL-1 Blue (Stratagene Cloning Systems, La Jolla, CA) cells are transformed with the ligation mixture and spread on culture plates. Bacteria containing the plasmid are identified by digesting DNA from individual colonies and analyzing the digestion product by gel electrophoresis. The sequence of the cloned fragment is confirmed by DNA sequencing.

Five µg of a plasmid containing the polynucleotide is co-transfected with 1.0 µg of a commercially available linearized baculovirus DNA ("BaculoGold™ baculovirus DNA", Pharmingen, San Diego, CA), using the lipofection method described by Felgner et al., Proc. Natl. Acad. Sci. USA 84:7413-7417 (1987). One µg of BaculoGold™ virus DNA and 5 µg of the plasmid are mixed in a sterile well of a microtiter plate containing 50 µl of serum-free Grace's medium (Life Technologies Inc., Gaithersburg, MD). Afterwards, 10 µl Lipofectin plus 90 µl Grace's medium are added, mixed and incubated for 15 minutes at room temperature. Then the transfection mixture is added drop-wise to Sf9 insect cells (ATCC CRL 1711) seeded in a 35 mm

tissue culture plate with 1 ml Grace's medium without serum. The plate is then incubated for 5 hours at 27° C. The transfection solution is then removed from the plate and 1 ml of Grace's insect medium supplemented with 10% fetal calf serum is added. Cultivation is then continued at 27° C for four days.

5 After four days the supernatant is collected and a plaque assay is performed, as described by Summers and Smith, *supra*. An agarose gel with "Blue Gal" (Life Technologies Inc., Gaithersburg) is used to allow easy identification and isolation of gal-expressing clones, which produce blue-stained plaques. (A detailed description of a "plaque assay" of this type can also be found in the user's guide for insect cell culture and baculovirology distributed by Life Technologies Inc., Gaithersburg, page 9-10.)  
10 After appropriate incubation, blue stained plaques are picked with the tip of a micropipettor (e.g., Eppendorf). The agar containing the recombinant viruses is then resuspended in a microcentrifuge tube containing 200 µl of Grace's medium and the suspension containing the recombinant baculovirus is used to infect Sf9 cells seeded in  
15 35 mm dishes. Four days later the supernatants of these culture dishes are harvested and then they are stored at 4° C.

To verify the expression of the polypeptide, Sf9 cells are grown in Grace's medium supplemented with 10% heat-inactivated FBS. The cells are infected with the recombinant baculovirus containing the polynucleotide at a multiplicity of infection ("MOI") of about 2. If radiolabeled proteins are desired, 6 hours later the medium is removed and is replaced with SF900 II medium minus methionine and cysteine (available from Life Technologies Inc., Rockville, MD). After 42 hours, 5 µCi of <sup>35</sup>S-methionine and 5 µCi <sup>35</sup>S-cysteine (available from Amersham) are added. The cells are further incubated for 16 hours and then are harvested by centrifugation. The proteins 20 in the supernatant as well as the intracellular proteins are analyzed by SDS-PAGE followed by autoradiography (if radiolabeled).

Microsequencing of the amino acid sequence of the amino terminus of purified protein may be used to determine the amino terminal sequence of the produced protein.

30 **Example 8: Expression of a Polypeptide in Mammalian Cells**

The polypeptide of the present invention can be expressed in a mammalian cell. A typical mammalian expression vector contains a promoter element, which mediates

the initiation of transcription of mRNA, a protein coding sequence, and signals required for the termination of transcription and polyadenylation of the transcript. Additional elements include enhancers, Kozak sequences and intervening sequences flanked by donor and acceptor sites for RNA splicing. Highly efficient transcription is achieved  
5 with the early and late promoters from SV40, the long terminal repeats (LTRs) from Retroviruses, e.g., RSV, HTLV, HIVI and the early promoter of the cytomegalovirus (CMV). However, cellular elements can also be used (e.g., the human actin promoter).

Suitable expression vectors for use in practicing the present invention include, for example, vectors such as pSVL and pMSG (Pharmacia, Uppsala, Sweden),  
10 pRSVcat (ATCC 37152), pSV2dhfr (ATCC 37146), pBC12MI (ATCC 67109), pCMVSport 2.0, and pCMVSport 3.0. Mammalian host cells that could be used include, human Hela, 293, H9 and Jurkat cells, mouse NIH3T3 and C127 cells, Cos 1, Cos 7 and CV1, quail QC1-3 cells, mouse L cells and Chinese hamster ovary (CHO) cells.

15 Alternatively, the polypeptide can be expressed in stable cell lines containing the polynucleotide integrated into a chromosome. The co-transfection with a selectable marker such as dhfr, gpt, neomycin, hygromycin allows the identification and isolation of the transfected cells.

The transfected gene can also be amplified to express large amounts of the  
20 encoded protein. The DHFR (dihydrofolate reductase) marker is useful in developing cell lines that carry several hundred or even several thousand copies of the gene of interest. (See, e.g., Alt, F. W., et al., J. Biol. Chem. 253:1357-1370 (1978); Hamlin,  
J. L. and Ma, C., Biochem. et Biophys. Acta, 1097:107-143 (1990); Page, M. J. and Sydenham, M. A., Biotechnology 9:64-68 (1991).) Another useful selection marker is  
25 the enzyme glutamine synthase (GS) (Murphy et al., Biochem J. 227:277-279 (1991); Bebbington et al., Bio/Technology 10:169-175 (1992). Using these markers, the mammalian cells are grown in selective medium and the cells with the highest resistance are selected. These cell lines contain the amplified gene(s) integrated into a chromosome. Chinese hamster ovary (CHO) and NSO cells are often used for the  
30 production of proteins.

Derivatives of the plasmid pSV2-dhfr (ATCC Accession No. 37146), the expression vectors pC4 (ATCC Accession No. 209646) and pC6 (ATCC Accession No. 209647) contain the strong promoter (LTR) of the Rous Sarcoma Virus (Cullen et al., Molecular and Cellular Biology, 438-447 (March, 1985)) plus a fragment of the  
35 CMV-enhancer (Boshart et al., Cell 41:521-530 (1985).) Multiple cloning sites, e.g., with the restriction enzyme cleavage sites BamHI, XbaI and Asp718, facilitate the cloning of the gene of interest. The vectors also contain the 3' intron, the

polyadenylation and termination signal of the rat preproinsulin gene, and the mouse DHFR gene under control of the SV40 early promoter.

Specifically, the plasmid pC6, for example, is digested with appropriate restriction enzymes and then dephosphorylated using calf intestinal phosphates by 5 procedures known in the art. The vector is then isolated from a 1% agarose gel.

A polynucleotide of the present invention is amplified according to the protocol outlined in Example 1. If the naturally occurring signal sequence is used to produce the secreted protein, the vector does not need a second signal peptide. Alternatively, if the naturally occurring signal sequence is not used, the vector can be modified to include a 10 heterologous signal sequence. (See, e.g., WO 96/34891.)

The amplified fragment is isolated from a 1% agarose gel using a commercially available kit ("Geneclean," BIO 101 Inc., La Jolla, Ca.). The fragment then is digested with appropriate restriction enzymes and again purified on a 1% agarose gel.

15 The amplified fragment is then digested with the same restriction enzyme and purified on a 1% agarose gel. The isolated fragment and the dephosphorylated vector are then ligated with T4 DNA ligase. *E. coli* HB101 or XL-1 Blue cells are then transformed and bacteria are identified that contain the fragment inserted into plasmid pC6 using, for instance, restriction enzyme analysis.

Chinese hamster ovary cells lacking an active DHFR gene is used for 20 transfection. Five  $\mu$ g of the expression plasmid pC6 is cotransfected with 0.5  $\mu$ g of the plasmid pSVneo using lipofectin (Felgner et al., *supra*). The plasmid pSV2-neo contains a dominant selectable marker, the *neo* gene from Tn5 encoding an enzyme that confers resistance to a group of antibiotics including G418. The cells are seeded in alpha minus MEM supplemented with 1 mg/ml G418. After 2 days, the cells are 25 trypsinized and seeded in hybridoma cloning plates (Greiner, Germany) in alpha minus MEM supplemented with 10, 25, or 50 ng/ml of methotrexate plus 1 mg/ml G418. After about 10-14 days single clones are trypsinized and then seeded in 6-well petri dishes or 10 ml flasks using different concentrations of methotrexate (50 nM, 100 nM, 200 nM, 400 nM, 800 nM). Clones growing at the highest concentrations of 30 methotrexate are then transferred to new 6-well plates containing even higher concentrations of methotrexate (1  $\mu$ M, 2  $\mu$ M, 5  $\mu$ M, 10 mM, 20 mM). The same procedure is repeated until clones are obtained which grow at a concentration of 100 - 200  $\mu$ M. Expression of the desired gene product is analyzed, for instance, by SDS-PAGE and Western blot or by reversed phase HPLC analysis.

35

**Example 9: Protein Fusions**

The polypeptides of the present invention are preferably fused to other proteins. These fusion proteins can be used for a variety of applications. For example, fusion of the present polypeptides to His-tag, HA-tag, protein A, IgG domains, and maltose binding protein facilitates purification. (See Example 5; see also EP A 394,827; 5 Traunecker, et al., Nature 331:84-86 (1988).) Similarly, fusion to IgG-1, IgG-3, and albumin increases the halflife time in vivo. Nuclear localization signals fused to the polypeptides of the present invention can target the protein to a specific subcellular localization, while covalent heterodimer or homodimers can increase or decrease the activity of a fusion protein. Fusion proteins can also create chimeric molecules having 10 more than one function. Finally, fusion proteins can increase solubility and/or stability of the fused protein compared to the non-fused protein. All of the types of fusion proteins described above can be made by modifying the following protocol, which outlines the fusion of a polypeptide to an IgG molecule, or the protocol described in Example 5.

15 Briefly, the human Fc portion of the IgG molecule can be PCR amplified, using primers that span the 5' and 3' ends of the sequence described below. These primers also should have convenient restriction enzyme sites that will facilitate cloning into an expression vector, preferably a mammalian expression vector.

20 For example, if pC4 (Accession No. 209646) is used, the human Fc portion can be ligated into the BamHI cloning site. Note that the 3' BamHI site should be destroyed. Next, the vector containing the human Fc portion is re-restricted with BamHI, linearizing the vector, and a polynucleotide of the present invention, isolated by the PCR protocol described in Example 1, is ligated into this BamHI site. Note that the polynucleotide is cloned without a stop codon, otherwise a fusion protein will not 25 be produced.

If the naturally occurring signal sequence is used to produce the secreted protein, pC4 does not need a second signal peptide. Alternatively, if the naturally occurring signal sequence is not used, the vector can be modified to include a heterologous signal sequence. (See, e.g., WO 96/34891.)

30

Human IgG Fc region:

GGGATCCGGAGCCAAATCTTCTGACAAAACTCACACATGCCACCCTGCC  
CAGCACCTGAATTGAGGGTGCACCGTCAGTCTTCTCTTCCCCCAAAACC  
CAAGGACACCCCTCATGATCTCCGGACTCCTGAGGTACATGCGTGGTGG  
35 GGACGTAAGCCACGAAGACCCCTGAGGTCAAGTTCAACTGGTACGTGGACG  
GCGTGGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAGTACAAC  
AGCACGTACCGTGTGGTCAGCGTCCTCACCGTCTGCACCAAGGACTGGCTG

AATGGCAAGGAGTACAAGTGCAAGGTCTCCAACAAAGCCCTCCCAACCCCC  
ATCGAGAAAACCATCTCCAAAGCCAAAGGGCAGCCCCGAGAACCCACAGGT  
GTACACCCTGCCCATCCGGATGAGCTGACCAAGAACCAAGGTAGCCT  
GACCTGCCTGGTCAAAGGCTTCTATCCAAGCGACATGCCGTGGAGTGGGA  
5 GAGCAATGGGCAGCCGGAGAACAACTACAAGACCACGCCTCCCGTGTGG  
ACTCCGACGGCTCCTCTTCCTCTACAGCAAGCTCACCGTGGACAAGAGCA  
GGTGGCAGCAGGGAACGTCTCTCATGCTCCGTATGCATGAGGCTCTGC  
ACAACCACTACACGCAGAAGAGCCTCTCCGTCTCCGGTAAATGAGTGC  
GACGGCCGCGACTCTAGAGGAT (SEQ ID NO:1)

10

**Example 10: Production of an Antibody from a Polypeptide**

The antibodies of the present invention can be prepared by a variety of methods. (See, Current Protocols, Chapter 2.) For example, cells expressing a polypeptide of the present invention is administered to an animal to induce the production of sera containing polyclonal antibodies. In a preferred method, a preparation of the secreted protein is prepared and purified to render it substantially free of natural contaminants. Such a preparation is then introduced into an animal in order to produce polyclonal antisera of greater specific activity.

In the most preferred method, the antibodies of the present invention are monoclonal antibodies (or protein binding fragments thereof). Such monoclonal antibodies can be prepared using hybridoma technology. (Köhler et al., *Nature* 256:495 (1975); Köhler et al., *Eur. J. Immunol.* 6:511 (1976); Köhler et al., *Eur. J. Immunol.* 6:292 (1976); Hammerling et al., in: *Monoclonal Antibodies and T-Cell Hybridomas*, Elsevier, N.Y., pp. 563-681 (1981).) In general, such procedures involve immunizing an animal (preferably a mouse) with polypeptide or, more preferably, with a secreted polypeptide-expressing cell. Such cells may be cultured in any suitable tissue culture medium; however, it is preferable to culture cells in Earle's modified Eagle's medium supplemented with 10% fetal bovine serum (inactivated at about 56°C), and supplemented with about 10 g/l of nonessential amino acids, about 30 1,000 U/ml of penicillin, and about 100 µg/ml of streptomycin.

The splenocytes of such mice are extracted and fused with a suitable myeloma cell line. Any suitable myeloma cell line may be employed in accordance with the present invention; however, it is preferable to employ the parent myeloma cell line (SP2O), available from the ATCC. After fusion, the resulting hybridoma cells are selectively maintained in HAT medium, and then cloned by limiting dilution as described by Wands et al. (*Gastroenterology* 80:225-232 (1981).) The hybridoma cells

obtained through such a selection are then assayed to identify clones which secrete antibodies capable of binding the polypeptide.

Alternatively, additional antibodies capable of binding to the polypeptide can be produced in a two-step procedure using anti-idiotypic antibodies. Such a method makes use of the fact that antibodies are themselves antigens, and therefore, it is possible to obtain an antibody which binds to a second antibody. In accordance with this method, protein specific antibodies are used to immunize an animal, preferably a mouse. The splenocytes of such an animal are then used to produce hybridoma cells, and the hybridoma cells are screened to identify clones which produce an antibody whose ability to bind to the protein-specific antibody can be blocked by the polypeptide. Such antibodies comprise anti-idiotypic antibodies to the protein-specific antibody and can be used to immunize an animal to induce formation of further protein-specific antibodies.

It will be appreciated that Fab and F(ab')2 and other fragments of the antibodies of the present invention may be used according to the methods disclosed herein. Such fragments are typically produced by proteolytic cleavage, using enzymes such as papain (to produce Fab fragments) or pepsin (to produce F(ab')2 fragments). Alternatively, secreted protein-binding fragments can be produced through the application of recombinant DNA technology or through synthetic chemistry.

For in vivo use of antibodies in humans, it may be preferable to use "humanized" chimeric monoclonal antibodies. Such antibodies can be produced using genetic constructs derived from hybridoma cells producing the monoclonal antibodies described above. Methods for producing chimeric antibodies are known in the art. (See, for review, Morrison, Science 229:1202 (1985); Oi et al., BioTechniques 4:214 (1986); Cabilly et al., U.S. Patent No. 4,816,567; Taniguchi et al., EP 171496; Morrison et al., EP 173494; Neuberger et al., WO 8601533; Robinson et al., WO 8702671; Boulianne et al., Nature 312:643 (1984); Neuberger et al., Nature 314:268 (1985).)

**30      Example 11: Production Of Secreted Protein For High-Throughput Screening Assays**

The following protocol produces a supernatant containing a polypeptide to be tested. This supernatant can then be used in the Screening Assays described in Examples 13-20.

35      First, dilute Poly-D-Lysine (644 587 Boehringer-Mannheim) stock solution (1mg/ml in PBS) 1:20 in PBS (w/o calcium or magnesium 17-516F Biowhittaker) for a working solution of 50ug/ml. Add 200 ul of this solution to each well (24 well plates)

and incubate at RT for 20 minutes. Be sure to distribute the solution over each well (note: a 12-channel pipetter may be used with tips on every other channel). Aspirate off the Poly-D-Lysine solution and rinse with 1ml PBS (Phosphate Buffered Saline). The PBS should remain in the well until just prior to plating the cells and plates may be poly-lysine coated in advance for up to two weeks.

5 Plate 293T cells (do not carry cells past P+20) at  $2 \times 10^5$  cells/well in .5ml DMEM(Dulbecco's Modified Eagle Medium)(with 4.5 G/L glucose and L-glutamine (12-604F Biowhittaker))/10% heat inactivated FBS(14-503F Biowhittaker)/1x Penstrep(17-602E Biowhittaker). Let the cells grow overnight.

10 The next day, mix together in a sterile solution basin: 300 ul Lipofectamine (18324-012 Gibco/BRL) and 5ml Optimem I (31985070 Gibco/BRL)/96-well plate. With a small volume multi-channel pipetter, aliquot approximately 2ug of an expression vector containing a polynucleotide insert, produced by the methods described in Examples 8 or 9, into an appropriately labeled 96-well round bottom plate. With a 15 multi-channel pipetter, add 50ul of the Lipofectamine/Optimem I mixture to each well. Pipette up and down gently to mix. Incubate at RT 15-45 minutes. After about 20 minutes, use a multi-channel pipetter to add 150ul Optimem I to each well. As a control, one plate of vector DNA lacking an insert should be transfected with each set of transfections.

20 Preferably, the transfection should be performed by tag-teaming the following tasks. By tag-teaming, hands on time is cut in half, and the cells do not spend too much time on PBS. First, person A aspirates off the media from four 24-well plates of cells, and then person B rinses each well with .5-1ml PBS. Person A then aspirates off PBS rinse, and person B, using a12-channel pipetter with tips on every other channel, 25 adds the 200ul of DNA/Lipofectamine/Optimem I complex to the odd wells first, then to the even wells, to each row on the 24-well plates. Incubate at 37°C for 6 hours.

While cells are incubating, prepare appropriate media, either 1%BSA in DMEM with 1x penstrep, or CHO-5 media (116.6 mg/L of CaCl<sub>2</sub> (anhyd); 0.00130 mg/L CuSO<sub>4</sub>·5H<sub>2</sub>O; 0.050 mg/L of Fe(NO<sub>3</sub>)<sub>3</sub>·9H<sub>2</sub>O; 0.417 mg/L of FeSO<sub>4</sub>·7H<sub>2</sub>O; 311.80 30 mg/L of Kcl; 28.64 mg/L of MgCl<sub>2</sub>; 48.84 mg/L of MgSO<sub>4</sub>; 6995.50 mg/L of NaCl; 2400.0 mg/L of NaHCO<sub>3</sub>; 62.50 mg/L of NaH<sub>2</sub>PO<sub>4</sub>·H<sub>2</sub>O; 71.02 mg/L of Na<sub>2</sub>HPO<sub>4</sub>; .4320 mg/L of ZnSO<sub>4</sub>·7H<sub>2</sub>O; .002 mg/L of Arachidonic Acid ; 1.022 mg/L of Cholesterol; .070 mg/L of DL-alpha-Tocopherol-Acetate; 0.0520 mg/L of Linoleic Acid; 0.010 mg/L of Linolenic Acid; 0.010 mg/L of Myristic Acid; 0.010 mg/L of Oleic 35 Acid; 0.010 mg/L of Palmitric Acid; 0.010 mg/L of Palmitic Acid; 100 mg/L of Pluronic F-68; 0.010 mg/L of Stearic Acid; 2.20 mg/L of Tween 80; 4551 mg/L of D-

Glucose; 130.85 mg/ml of L-Alanine; 147.50 mg/ml of L-Arginine-HCL; 7.50 mg/ml of L-Asparagine-H<sub>2</sub>O; 6.65 mg/ml of L-Aspartic Acid; 29.56 mg/ml of L-Cystine-2HCL-H<sub>2</sub>O; 31.29 mg/ml of L-Cystine-2HCL; 7.35 mg/ml of L-Glutamic Acid; 365.0 mg/ml of L-Glutamine; 18.75 mg/ml of Glycine; 52.48 mg/ml of L-Histidine-HCL-H<sub>2</sub>O; 106.97 mg/ml of L-Isoleucine; 111.45 mg/ml of L-Leucine; 163.75 mg/ml of L-Lysine HCL; 32.34 mg/ml of L-Methionine; 68.48 mg/ml of L-Phenylalanine; 40.0 mg/ml of L-Proline; 26.25 mg/ml of L-Serine; 101.05 mg/ml of L-Threonine; 19.22 mg/ml of L-Tryptophan; 91.79 mg/ml of L-Tyrosine-2Na-2H<sub>2</sub>O; 99.65 mg/ml of L-Valine; 0.0035 mg/L of Biotin; 3.24 mg/L of D-Ca Pantothenate; 11.78 mg/L of Choline Chloride; 4.65 mg/L of Folic Acid; 15.60 mg/L of i-Inositol; 3.02 mg/L of Niacinamide; 3.00 mg/L of Pyridoxal HCL; 0.031 mg/L of Pyridoxine HCL; 0.319 mg/L of Riboflavin; 3.17 mg/L of Thiamine HCL; 0.365 mg/L of Thymidine; and 0.680 mg/L of Vitamin B<sub>12</sub>; 25 mM of HEPES Buffer; 2.39 mg/L of Na Hypoxanthine; 0.105 mg/L of Lipoic Acid; 0.081 mg/L of Sodium Putrescine-2HCL; 55.0 mg/L of Sodium Pyruvate; 0.0067 mg/L of Sodium Selenite; 20uM of Ethanolamine; 0.122 mg/L of Ferric Citrate; 41.70 mg/L of Methyl-B-Cyclodextrin complexed with Linoleic Acid; 33.33 mg/L of Methyl-B-Cyclodextrin complexed with Oleic Acid; and 10 mg/L of Methyl-B-Cyclodextrin complexed with Retinal) with 2mm glutamine and 1x penstrep. (BSA (81-068-3 Bayer) 100gm dissolved in 1L DMEM for a 10% BSA stock solution). Filter the media and collect 50 ul for endotoxin assay in 15ml polystyrene conical.

The transfection reaction is terminated, preferably by tag-teaming, at the end of the incubation period. Person A aspirates off the transfection media, while person B adds 1.5ml appropriate media to each well. Incubate at 37°C for 45 or 72 hours depending on the media used: 1%BSA for 45 hours or CHO-5 for 72 hours.

On day four, using a 300ul multichannel pipetter, aliquot 600ul in one 1ml deep well plate and the remaining supernatant into a 2ml deep well. The supernatants from each well can then be used in the assays described in Examples 13-20.

It is specifically understood that when activity is obtained in any of the assays described below using a supernatant, the activity originates from either the polypeptide directly (e.g., as a secreted protein) or by the polypeptide inducing expression of other proteins, which are then secreted into the supernatant. Thus, the invention further provides a method of identifying the protein in the supernatant characterized by an activity in a particular assay.

**Example 12: Construction of GAS Reporter Construct**

One signal transduction pathway involved in the differentiation and proliferation of cells is called the Jaks-STATs pathway. Activated proteins in the Jaks-STATs pathway bind to gamma activation site "GAS" elements or interferon-sensitive responsive element ("ISRE"), located in the promoter of many genes. The binding of a protein to these elements alter the expression of the associated gene.

GAS and ISRE elements are recognized by a class of transcription factors called Signal Transducers and Activators of Transcription, or "STATs." There are six members of the STATs family. Stat1 and Stat3 are present in many cell types, as is Stat2 (as response to IFN-alpha is widespread). Stat4 is more restricted and is not in many cell types though it has been found in T helper class I, cells after treatment with IL-12. Stat5 was originally called mammary growth factor, but has been found at higher concentrations in other cells including myeloid cells. It can be activated in tissue culture cells by many cytokines.

The STATs are activated to translocate from the cytoplasm to the nucleus upon tyrosine phosphorylation by a set of kinases known as the Janus Kinase ("Jaks") family. Jaks represent a distinct family of soluble tyrosine kinases and include Tyk2, Jak1, Jak2, and Jak3. These kinases display significant sequence similarity and are generally catalytically inactive in resting cells.

The Jaks are activated by a wide range of receptors summarized in the Table below. (Adapted from review by Schidler and Darnell, Ann. Rev. Biochem. 64:621-51 (1995).) A cytokine receptor family, capable of activating Jaks, is divided into two groups: (a) Class 1 includes receptors for IL-2, IL-3, IL-4, IL-6, IL-7, IL-9, IL-11, IL-12, IL-15, Epo, PRL, GH, G-CSF, GM-CSF, LIF, CNTF, and thrombopoietin; and (b) Class 2 includes IFN- $\alpha$ , IFN- $\gamma$ , and IL-10. The Class 1 receptors share a conserved cysteine motif (a set of four conserved cysteines and one tryptophan) and a WSXWS motif (a membrane proximal region encoding Trp-Ser-Xxx-Trp-Ser (SEQ ID NO:2)).

Thus, on binding of a ligand to a receptor, Jaks are activated, which in turn activate STATs, which then translocate and bind to GAS elements. This entire process is encompassed in the Jaks-STATs signal transduction pathway.

Therefore, activation of the Jaks-STATs pathway, reflected by the binding of the GAS or the ISRE element, can be used to indicate proteins involved in the proliferation and differentiation of cells. For example, growth factors and cytokines are known to activate the Jaks-STATs pathway. (See Table below.) Thus, by using GAS elements linked to reporter molecules, activators of the Jaks-STATs pathway can be identified.

		<u>JAKs</u>		<u>STATS</u>	<u>GAS(elements) or ISRE</u>	
	<u>Ligand</u>	<u>tyk2</u>	<u>Jak1</u>	<u>Jak2</u>	<u>Jak3</u>	
<u>IFN family</u>						
5	IFN-a/B	+	+	-	-	1,2,3 ISRE
	IFN-g		+	+	-	1 GAS (IRF1>Lys6>IFP)
	Il-10	+	?	?	-	1,3
<u>gp130 family</u>						
10	IL-6 (Pleiotrophic)	+	+	+	?	1,3 GAS (IRF1>Lys6>IFP)
	Il-11(Pleiotrophic)	?	+	?	?	1,3
	OnM(Pleiotrophic)	?	+	+	?	1,3
	LIF(Pleiotrophic)	?	+	+	?	1,3
15	CNTF(Pleiotrophic)	-/+	+	+	?	1,3
	G-CSF(Pleiotrophic)	?	+	?	?	1,3
	IL-12(Pleiotrophic)	+	-	+	+	1,3
<u>g-C family</u>						
20	IL-2 (lymphocytes)	-	+	-	+	1,3,5 GAS
	IL-4 (lymph/myeloid)	-	+	-	+	6 GAS (IRF1 = IFP >>Ly6)(IgH)
	IL-7 (lymphocytes)	-	+	-	+	5 GAS
	IL-9 (lymphocytes)	-	+	-	+	5 GAS
	IL-13 (lymphocyte)	-	+	?	?	6 GAS
	IL-15	?	+	?	+	5 GAS
<u>gp140 family</u>						
25	IL-3 (myeloid)	-	-	+	-	5 GAS (IRF1>IFP>>Ly6)
	IL-5 (myeloid)	-	-	+	-	5 GAS
	GM-CSF (myeloid)	-	-	+	-	5 GAS
<u>Growth hormone family</u>						
30	GH	?	-	+	-	5
	PRL	?	+/-	+	-	1,3,5
	EPO	?	-	+	-	5 GAS(B-CAS>IRF1=IFP>>Ly6)
<u>Receptor Tyrosine Kinases</u>						
35	EGF	?	+	+	-	1,3 GAS (IRF1)
	PDGF	?	+	+	-	1,3
	CSF-1	?	+	+	-	1,3 GAS (not IRF1)
40						

To construct a synthetic GAS containing promoter element, which is used in the Biological Assays described in Examples 13-14, a PCR based strategy is employed to generate a GAS-SV40 promoter sequence. The 5' primer contains four tandem copies of the GAS binding site found in the IRF1 promoter and previously demonstrated to bind STATs upon induction with a range of cytokines (Rothman et al., *Immunity* 1:457-468 (1994).), although other GAS or ISRE elements can be used instead. The 5' primer also contains 18bp of sequence complementary to the SV40 early promoter sequence and is flanked with an XhoI site. The sequence of the 5' primer is:

5':GCGCCTCGAGATTCCCCGAAATCTAGATTCCCCGAAATGATTCCCCG  
10 AAATGATTCCCCGAAATATCTGCCATCTCAATTAG:3' (SEQ ID NO:3)

The downstream primer is complementary to the SV40 promoter and is flanked with a Hind III site: 5':GCGGCAAGCTTTGCAAAGCCTAGGC:3' (SEQ ID NO:4)

PCR amplification is performed using the SV40 promoter template present in the B-gal:promoter plasmid obtained from Clontech. The resulting PCR fragment is digested with XhoI/Hind III and subcloned into BLSK2-. (Stratagene.) Sequencing with forward and reverse primers confirms that the insert contains the following sequence:

5':CTCGAGATTCCCCGAAATCTAGATTCCCCGAAATGATTCCCCGAAATG  
20 ATTCCCCGAAATATCTGCCATCTCAATTAGTCAGCAACCATAGTCCCC  
CTAACTCCGGCCCCATCCCGCCCTAACTCCGCCCAGTTCCGCCATTCTCCGC  
CCCATGGCTGACTATTTTTTTTATGCAGAGGCCGAGGCCGCGCTCGGC  
CTCTGAGCTATTCCAGAGTAGGTGAGGAGGGCTTTTGGAGGCCTAGGCTTTT  
TGCAAAAAGCTT:3' (SEQ ID NO:5)

With this GAS promoter element linked to the SV40 promoter, a GAS:SEAP2 reporter construct is next engineered. Here, the reporter molecule is a secreted alkaline phosphatase, or "SEAP." Clearly, however, any reporter molecule can be instead of SEAP, in this or in any of the other Examples. Well known reporter molecules that can be used instead of SEAP include chloramphenicol acetyltransferase (CAT), luciferase, alkaline phosphatase, B-galactosidase, green fluorescent protein (GFP), or any protein detectable by an antibody.

The above sequence confirmed synthetic GAS-SV40 promoter element is subcloned into the pSEAP-Promoter vector obtained from Clontech using HindIII and XhoI, effectively replacing the SV40 promoter with the amplified GAS:SV40 promoter element, to create the GAS-SEAP vector. However, this vector does not contain a neomycin resistance gene, and therefore, is not preferred for mammalian expression systems.

Thus, in order to generate mammalian stable cell lines expressing the GAS-SEAP reporter, the GAS-SEAP cassette is removed from the GAS-SEAP vector using SalI and NotI, and inserted into a backbone vector containing the neomycin resistance gene, such as pGFP-1 (Clontech), using these restriction sites in the multiple cloning site, to create the GAS-SEAP/Neo vector. Once this vector is transfected into mammalian cells, this vector can then be used as a reporter molecule for GAS binding as described in Examples 13-14.

Other constructs can be made using the above description and replacing GAS with a different promoter sequence. For example, construction of reporter molecules containing NFK-B and EGR promoter sequences are described in Examples 15 and 16. However, many other promoters can be substituted using the protocols described in these Examples. For instance, SRE, IL-2, NFAT, or Osteocalcin promoters can be substituted, alone or in combination (e.g., GAS/NF-KB/EGR, GAS/NF-KB, IL-2/NFAT, or NF-KB/GAS). Similarly, other cell lines can be used to test reporter construct activity, such as HELA (epithelial), HUVEC (endothelial), Reh (B-cell), Saos-2 (osteoblast), HUVAC (aortic), or Cardiomyocyte.

**Example 13: High-Throughput Screening Assay for T-cell Activity.**

The following protocol is used to assess T-cell activity by identifying factors, such as growth factors and cytokines, that may proliferate or differentiate T-cells. T-cell activity is assessed using the GAS/SEAP/Neo construct produced in Example 12. Thus, factors that increase SEAP activity indicate the ability to activate the Jaks-STATS signal transduction pathway. The T-cell used in this assay is Jurkat T-cells (ATCC Accession No. TIB-152), although Molt-3 cells (ATCC Accession No. CRL-1552) and Molt-4 cells (ATCC Accession No. CRL-1582) cells can also be used.

Jurkat T-cells are lymphoblastic CD4+ Th1 helper cells. In order to generate stable cell lines, approximately 2 million Jurkat cells are transfected with the GAS-SEAP/neo vector using DMRIE-C (Life Technologies)(transfection procedure described below). The transfected cells are seeded to a density of approximately 20,000 cells per well and transfectants resistant to 1 mg/ml gentamicin selected. Resistant colonies are expanded and then tested for their response to increasing concentrations of interferon gamma. The dose response of a selected clone is demonstrated.

Specifically, the following protocol will yield sufficient cells for 75 wells containing 200 ul of cells. Thus, it is either scaled up, or performed in multiple to generate sufficient cells for multiple 96 well plates. Jurkat cells are maintained in RPMI + 10% serum with 1%Pen-Strep. Combine 2.5 mls of OPTI-MEM (Life Technologies)

with 10 ug of plasmid DNA in a T25 flask. Add 2.5 ml OPTI-MEM containing 50 ul of DMRIE-C and incubate at room temperature for 15-45 mins.

During the incubation period, count cell concentration, spin down the required number of cells ( $10^7$  per transfection), and resuspend in OPTI-MEM to a final concentration of  $10^7$  cells/ml. Then add 1ml of  $1 \times 10^7$  cells in OPTI-MEM to T25 flask and incubate at 37°C for 6 hrs. After the incubation, add 10 ml of RPMI + 15% serum.

The Jurkat:GAS-SEAP stable reporter lines are maintained in RPMI + 10% serum, 1 mg/ml Genticin, and 1% Pen-Strep. These cells are treated with supernatants containing a polypeptide as produced by the protocol described in Example 11.

On the day of treatment with the supernatant, the cells should be washed and resuspended in fresh RPMI + 10% serum to a density of 500,000 cells per ml. The exact number of cells required will depend on the number of supernatants being screened. For one 96 well plate, approximately 10 million cells (for 10 plates, 100 million cells) are required.

Transfer the cells to a triangular reservoir boat, in order to dispense the cells into a 96 well dish, using a 12 channel pipette. Using a 12 channel pipette, transfer 200 ul of cells into each well (therefore adding 100,000 cells per well).

After all the plates have been seeded, 50 ul of the supernatants are transferred directly from the 96 well plate containing the supernatants into each well using a 12 channel pipette. In addition, a dose of exogenous interferon gamma (0.1, 1.0, 10 ng) is added to wells H9, H10, and H11 to serve as additional positive controls for the assay.

The 96 well dishes containing Jurkat cells treated with supernatants are placed in an incubator for 48 hrs (note: this time is variable between 48-72 hrs). 35 ul samples from each well are then transferred to an opaque 96 well plate using a 12 channel pipette. The opaque plates should be covered (using sellophene covers) and stored at -20°C until SEAP assays are performed according to Example 17. The plates containing the remaining treated cells are placed at 4°C and serve as a source of material for repeating the assay on a specific well if desired.

As a positive control, 100 Unit/ml interferon gamma can be used which is known to activate Jurkat T cells. Over 30 fold induction is typically observed in the positive control wells.

**Example 14: High-Throughput Screening Assay Identifying Myeloid Activity**

The following protocol is used to assess myeloid activity by identifying factors, such as growth factors and cytokines, that may proliferate or differentiate myeloid cells. Myeloid cell activity is assessed using the GAS/SEAP/Neo construct produced in Example 12. Thus, factors that increase SEAP activity indicate the ability to activate the Jak-STATS signal transduction pathway. The myeloid cell used in this assay is U937, a pre-monocyte cell line, although TF-1, HL60, or KG1 can be used.

To transiently transfet U937 cells with the GAS/SEAP/Neo construct produced in Example 12, a DEAE-Dextran method (Kharbanda et. al., 1994, Cell Growth & Differentiation, 5:259-265) is used. First, harvest  $2 \times 10^7$  U937 cells and wash with PBS. The U937 cells are usually grown in RPMI 1640 medium containing 10% heat-inactivated fetal bovine serum (FBS) supplemented with 100 units/ml penicillin and 100 mg/ml streptomycin.

Next, suspend the cells in 1 ml of 20 mM Tris-HCl (pH 7.4) buffer containing 0.5 mg/ml DEAE-Dextran, 8 ug GAS-SEAP2 plasmid DNA, 140 mM NaCl, 5 mM KCl, 375 uM Na<sub>2</sub>HPO<sub>4</sub>.7H<sub>2</sub>O, 1 mM MgCl<sub>2</sub>, and 675 uM CaCl<sub>2</sub>. Incubate at 37°C for 45 min.

Wash the cells with RPMI 1640 medium containing 10% FBS and then resuspend in 10 ml complete medium and incubate at 37°C for 36 hr.

The GAS-SEAP/U937 stable cells are obtained by growing the cells in 400 ug/ml G418. The G418-free medium is used for routine growth but every one to two months, the cells should be re-grown in 400 ug/ml G418 for couple of passages.

These cells are tested by harvesting  $1 \times 10^8$  cells (this is enough for ten 96-well plates assay) and wash with PBS. Suspend the cells in 200 ml above described growth medium, with a final density of  $5 \times 10^5$  cells/ml. Plate 200 ul cells per well in the 96-well plate (or  $1 \times 10^5$  cells/well).

Add 50 ul of the supernatant prepared by the protocol described in Example 11. Incubate at 37°C for 48 to 72 hr. As a positive control, 100 Unit/ml interferon gamma can be used which is known to activate U937 cells. Over 30 fold induction is typically observed in the positive control wells. SEAP assay the supernatant according to the protocol described in Example 17.

#### Example 15: High-Throughput Screening Assay Identifying Neuronal Activity.

When cells undergo differentiation and proliferation, a group of genes are activated through many different signal transduction pathways. One of these genes, EGR1 (early growth response gene 1), is induced in various tissues and cell types upon

activation. The promoter of EGR1 is responsible for such induction. Using the EGR1 promoter linked to reporter molecules, activation of cells can be assessed.

Particularly, the following protocol is used to assess neuronal activity in PC12 cell lines. PC12 cells (rat phenochromocytoma cells) are known to proliferate and/or differentiate by activation with a number of mitogens, such as TPA (tetradecanoyl phorbol acetate), NGF (nerve growth factor), and EGF (epidermal growth factor). The EGR1 gene expression is activated during this treatment. Thus, by stably transfecting PC12 cells with a construct containing an EGR promoter linked to SEAP reporter, activation of PC12 cells can be assessed.

10 The EGR/SEAP reporter construct can be assembled by the following protocol. The EGR-1 promoter sequence (-633 to +1)(Sakamoto K et al., Oncogene 6:867-871 (1991)) can be PCR amplified from human genomic DNA using the following primers:

5' GCGCTCGAGGGATGACAGCGATAGAACCCGG -3' (SEQ ID NO:6)

5' GCGAAGCTTCGCGACTCCCCGGATCCGCCTC-3' (SEQ ID NO:7)

15 Using the GAS:SEAP/Neo vector produced in Example 12, EGR1 amplified product can then be inserted into this vector. Linearize the GAS:SEAP/Neo vector using restriction enzymes XhoI/HindIII, removing the GAS/SV40 stuffer. Restrict the EGR1 amplified product with these same enzymes. Ligate the vector and the EGR1 promoter.

20 To prepare 96 well-plates for cell culture, two mls of a coating solution (1:30 dilution of collagen type I (Upstate Biotech Inc. Cat#08-115) in 30% ethanol (filter sterilized)) is added per one 10 cm plate or 50 ml per well of the 96-well plate, and allowed to air dry for 2 hr.

25 PC12 cells are routinely grown in RPMI-1640 medium (Bio Whittaker) containing 10% horse serum (JRH BIOSCIENCES, Cat. # 12449-78P), 5% heat-inactivated fetal bovine serum (FBS) supplemented with 100 units/ml penicillin and 100 ug/ml streptomycin on a precoated 10 cm tissue culture dish. One to four split is done every three to four days. Cells are removed from the plates by scraping and resuspended with pipetting up and down for more than 15 times.

30 Transfect the EGR/SEAP/Neo construct into PC12 using the Lipofectamine protocol described in Example 11. EGR-SEAP/PC12 stable cells are obtained by growing the cells in 300 ug/ml G418. The G418-free medium is used for routine growth but every one to two months, the cells should be re-grown in 300 ug/ml G418 for couple of passages.

35 To assay for neuronal activity, a 10 cm plate with cells around 70 to 80% confluent is screened by removing the old medium. Wash the cells once with PBS

(Phosphate buffered saline). Then starve the cells in low serum medium (RPMI-1640 containing 1% horse serum and 0.5% FBS with antibiotics) overnight.

The next morning, remove the medium and wash the cells with PBS. Scrape off the cells from the plate, suspend the cells well in 2 ml low serum medium. Count  
5 the cell number and add more low serum medium to reach final cell density as  $5 \times 10^5$  cells/ml.

Add 200 ul of the cell suspension to each well of 96-well plate (equivalent to  
1x $10^5$  cells/well). Add 50 ul supernatant produced by Example 11, 37°C for 48 to 72 hr. As a positive control, a growth factor known to activate PC12 cells through EGR  
10 can be used, such as 50 ng/ul of Neuronal Growth Factor (NGF). Over fifty-fold induction of SEAP is typically seen in the positive control wells. SEAP assay the supernatant according to Example 17.

**Example 16: High-Throughput Screening Assay for T-cell Activity**

NF- $\kappa$ B (Nuclear Factor  $\kappa$ B) is a transcription factor activated by a wide variety of agents including the inflammatory cytokines IL-1 and TNF, CD30 and CD40, lymphotoxin-alpha and lymphotoxin-beta, by exposure to LPS or thrombin, and by expression of certain viral gene products. As a transcription factor, NF- $\kappa$ B regulates the expression of genes involved in immune cell activation, control of apoptosis (NF-  
20  $\kappa$ B appears to shield cells from apoptosis), B and T-cell development, anti-viral and antimicrobial responses, and multiple stress responses.

In non-stimulated conditions, NF-  $\kappa$ B is retained in the cytoplasm with I- $\kappa$ B (Inhibitor  $\kappa$ B). However, upon stimulation, I-  $\kappa$ B is phosphorylated and degraded, causing NF-  $\kappa$ B to shuttle to the nucleus, thereby activating transcription of target  
25 genes. Target genes activated by NF-  $\kappa$ B include IL-2, IL-6, GM-CSF, ICAM-1 and class 1 MHC.

Due to its central role and ability to respond to a range of stimuli, reporter constructs utilizing the NF- $\kappa$ B promoter element are used to screen the supernatants produced in Example 11. Activators or inhibitors of NF- $\kappa$ B would be useful in treating  
30 diseases. For example, inhibitors of NF- $\kappa$ B could be used to treat those diseases related to the acute or chronic activation of NF- $\kappa$ B, such as rheumatoid arthritis.

To construct a vector containing the NF- $\kappa$ B promoter element, a PCR based strategy is employed. The upstream primer contains four tandem copies of the NF- $\kappa$ B binding site (GGGGACTTCCC) (SEQ ID NO:8), 18 bp of sequence complementary to the 5' end of the SV40 early promoter sequence, and is flanked with an XhoI site:

5' 5':GCGGCCTCGAGGGGACTTCCCGGGACTTCCGGGACTTCCGGGAC  
TTCCATCCTGCCATCTCAATTAG:3' (SEQ ID NO:9)

The downstream primer is complementary to the 3' end of the SV40 promoter and is flanked with a Hind III site:

5':GCGGCAAGCTTTGCAAAGCCTAGGC:3' (SEQ ID NO:4)

10 PCR amplification is performed using the SV40 promoter template present in the pB-gal:promoter plasmid obtained from Clontech. The resulting PCR fragment is digested with XhoI and Hind III and subcloned into BLSK2-. (Stratagene) Sequencing with the T7 and T3 primers confirms the insert contains the following sequence:

15 5':CTCGAGGGGACTTCCCGGGACTTCCGGGACTTCCGGGACTTCC  
ATCTGCCATCTCAATTAGTCAGCAACCATACTCCGCCCTAACTCCGCCA  
TCCCGCCCTAACTCCGCCAGTTCCGCCATTCTCCGCCCATGGCTGACT  
AATTTTTTTATTATGCAGAGGCCGAGGCCCTCGGCCTTGAGCTATT  
20 CAGAAGTAGTGAGGAGGCTTTTGAGGCCTAGGCTTTGCAAAAGCTT:  
3' (SEQ ID NO:10)

Next, replace the SV40 minimal promoter element present in the pSEAP2- promoter plasmid (Clontech) with this NF- $\kappa$ B/SV40 fragment using XhoI and HindIII.

25 However, this vector does not contain a neomycin resistance gene, and therefore, is not preferred for mammalian expression systems.

In order to generate stable mammalian cell lines, the NF- $\kappa$ B/SV40/SEAP cassette is removed from the above NF- $\kappa$ B/SEAP vector using restriction enzymes SalI and NotI, and inserted into a vector containing neomycin resistance. Particularly, the 30 NF- $\kappa$ B/SV40/SEAP cassette was inserted into pGFP-1 (Clontech), replacing the GFP gene, after restricting pGFP-1 with SalI and NotI.

Once NF- $\kappa$ B/SV40/SEAP/Neo vector is created, stable Jurkat T-cells are created and maintained according to the protocol described in Example 13. Similarly, the method for assaying supernatants with these stable Jurkat T-cells is also described

in Example 13. As a positive control, exogenous TNF alpha (0.1, 1, 10 ng) is added to wells H9, H10, and H11, with a 5-10 fold activation typically observed.

**Example 17: Assay for SEAP Activity**

5        As a reporter molecule for the assays described in Examples 13-16, SEAP activity is assayed using the Tropix Phospho-light Kit (Cat. BP-400) according to the following general procedure. The Tropix Phospho-light Kit supplies the Dilution, Assay, and Reaction Buffers used below.

10      Prime a dispenser with the 2.5x Dilution Buffer and dispense 15 µl of 2.5x dilution buffer into Optiplates containing 35 µl of a supernatant. Seal the plates with a plastic sealer and incubate at 65°C for 30 min. Separate the Optiplates to avoid uneven heating.

15      Cool the samples to room temperature for 15 minutes. Empty the dispenser and prime with the Assay Buffer. Add 50 µl Assay Buffer and incubate at room temperature 5 min. Empty the dispenser and prime with the Reaction Buffer (see the table below). Add 50 µl Reaction Buffer and incubate at room temperature for 20 minutes. Since the intensity of the chemiluminescent signal is time dependent, and it takes about 10 minutes to read 5 plates on luminometer, one should treat 5 plates at each time and start the second set 10 minutes later.

20      Read the relative light unit in the luminometer. Set H12 as blank, and print the results. An increase in chemiluminescence indicates reporter activity.

**Reaction Buffer Formulation:**

# of plates	Rxn buffer diluent (ml)	CSPD (ml)
10	60	3
11	65	3.25
12	70	3.5
13	75	3.75
14	80	4
15	85	4.25
16	90	4.5
17	95	4.75
18	100	5
19	105	5.25
20	110	5.5
21	115	5.75
22	120	6
23	125	6.25
24	130	6.5
25	135	6.75
26	140	7
27	145	7.25

28	150	7.5
29	155	7.75
30	160	8
31	165	8.25
32	170	8.5
33	175	8.75
34	180	9
35	185	9.25
36	190	9.5
37	195	9.75
38	200	10
39	205	10.25
40	210	10.5
41	215	10.75
42	220	11
43	225	11.25
44	230	11.5
45	235	11.75
46	240	12
47	245	12.25
48	250	12.5
49	255	12.75
50	260	13

**Example 18: High-Throughput Screening Assay Identifying Changes in Small Molecule Concentration and Membrane Permeability**

Binding of a ligand to a receptor is known to alter intracellular levels of small molecules, such as calcium, potassium, sodium, and pH, as well as alter membrane potential. These alterations can be measured in an assay to identify supernatants which bind to receptors of a particular cell. Although the following protocol describes an assay for calcium, this protocol can easily be modified to detect changes in potassium, sodium, pH, membrane potential, or any other small molecule which is detectable by a fluorescent probe:

The following assay uses Fluorometric Imaging Plate Reader ("FLIPR") to measure changes in fluorescent molecules (Molecular Probes) that bind small molecules. Clearly, any fluorescent molecule detecting a small molecule can be used instead of the calcium fluorescent molecule, fluo-3, used here.

For adherent cells, seed the cells at 10,000 -20,000 cells/well in a Co-star black 96-well plate with clear bottom. The plate is incubated in a CO<sub>2</sub> incubator for 20 hours. The adherent cells are washed two times in Biotek washer with 200 ul of HBSS (Hank's Balanced Salt Solution) leaving 100 ul of buffer after the final wash.

A stock solution of 1 mg/ml fluo-3 is made in 10% pluronic acid DMSO. To load the cells with fluo-3, 50 ul of 12 ug/ml fluo-3 is added to each well. The plate is

incubated at 37°C in a CO<sub>2</sub> incubator for 60 min. The plate is washed four times in the Biotek washer with HBSS leaving 100 ul of buffer.

For non-adherent cells, the cells are spun down from culture media. Cells are re-suspended to 2-5x10<sup>6</sup> cells/ml with HBSS in a 50-ml conical tube. 4 ul of 1 mg/ml fluo-3 solution in 10% pluronic acid DMSO is added to each ml of cell suspension. The tube is then placed in a 37°C water bath for 30-60 min. The cells are washed twice with HBSS, resuspended to 1x10<sup>6</sup> cells/ml, and dispensed into a microplate, 100 ul/well. The plate is centrifuged at 1000 rpm for 5 min. The plate is then washed once in Denley CellWash with 200 ul, followed by an aspiration step to 100 ul final volume.

For a non-cell based assay, each well contains a fluorescent molecule, such as fluo-3. The supernatant is added to the well, and a change in fluorescence is detected.

To measure the fluorescence of intracellular calcium, the FLIPR is set for the following parameters: (1) System gain is 300-800 mW; (2) Exposure time is 0.4 second; (3) Camera F/stop is F/2; (4) Excitation is 488 nm; (5) Emission is 530 nm; and (6) Sample addition is 50 ul. Increased emission at 530 nm indicates an extracellular signaling event which has resulted in an increase in the intracellular Ca<sup>++</sup> concentration.

#### Example 19: High-Throughput Screening Assay Identifying Tyrosine Kinase Activity

The Protein Tyrosine Kinases (PTK) represent a diverse group of transmembrane and cytoplasmic kinases. Within the Receptor Protein Tyrosine Kinase (RPTK) group are receptors for a range of mitogenic and metabolic growth factors including the PDGF, FGF, EGF, NGF, HGF and Insulin receptor subfamilies. In addition there are a large family of RPTKs for which the corresponding ligand is unknown. Ligands for RPTKs include mainly secreted small proteins, but also membrane-bound and extracellular matrix proteins.

Activation of RPTK by ligands involves ligand-mediated receptor dimerization, resulting in transphosphorylation of the receptor subunits and activation of the cytoplasmic tyrosine kinases. The cytoplasmic tyrosine kinases include receptor associated tyrosine kinases of the src-family (e.g., src, yes, lck, lyn, fyn) and non-receptor linked and cytosolic protein tyrosine kinases, such as the Jak family, members of which mediate signal transduction triggered by the cytokine superfamily of receptors (e.g., the Interleukins, Interferons, GM-CSF, and Leptin).

Because of the wide range of known factors capable of stimulating tyrosine kinase activity, the identification of novel human secreted proteins capable of activating

tyrosine kinase signal transduction pathways are of interest. Therefore, the following protocol is designed to identify those novel human secreted proteins capable of activating the tyrosine kinase signal transduction pathways.

Seed target cells (e.g., primary keratinocytes) at a density of approximately 5 25,000 cells per well in a 96 well Loprodynne Silent Screen Plates purchased from Nalge Nunc (Naperville, IL). The plates are sterilized with two 30 minute rinses with 100% ethanol, rinsed with water and dried overnight. Some plates are coated for 2 hr with 100 ml of cell culture grade type I collagen (50 mg/ml), gelatin (2%) or polylysine (50 mg/ml), all of which can be purchased from Sigma Chemicals (St. Louis, MO) or 10 10% Matrigel purchased from Becton Dickinson (Bedford, MA), or calf serum, rinsed with PBS and stored at 4°C. Cell growth on these plates is assayed by seeding 5,000 cells/well in growth medium and indirect quantitation of cell number through use of alamarBlue as described by the manufacturer Alamar Biosciences, Inc. (Sacramento, CA) after 48 hr. Falcon plate covers #3071 from Becton Dickinson (Bedford, MA) are 15 used to cover the Loprodynne Silent Screen Plates. Falcon Microtest III cell culture plates can also be used in some proliferation experiments.

To prepare extracts, A431 cells are seeded onto the nylon membranes of 20 Loprodynne plates (20,000/200ml/well) and cultured overnight in complete medium. Cells are quiesced by incubation in serum-free basal medium for 24 hr. After 5-20 minutes treatment with EGF (60ng/ml) or 50 ul of the supernatant produced in Example 11, the medium was removed and 100 ml of extraction buffer ((20 mM HEPES pH 7.5, 0.15 M NaCl, 1% Triton X-100, 0.1% SDS, 2 mM Na3VO4, 2 mM Na4P2O7 and a cocktail of protease inhibitors (# 1836170) obtained from Boehringer Mannheim (Indianapolis, IN) is added to each well and the plate is shaken on a rotating shaker for 25 5 minutes at 4°C. The plate is then placed in a vacuum transfer manifold and the extract filtered through the 0.45 mm membrane bottoms of each well using house vacuum. Extracts are collected in a 96-well catch/assay plate in the bottom of the vacuum manifold and immediately placed on ice. To obtain extracts clarified by centrifugation, the content of each well, after detergent solubilization for 5 minutes, is removed and 30 centrifuged for 15 minutes at 4°C at 16,000 x g.

Test the filtered extracts for levels of tyrosine kinase activity. Although many methods of detecting tyrosine kinase activity are known, one method is described here.

Generally, the tyrosine kinase activity of a supernatant is evaluated by 35 determining its ability to phosphorylate a tyrosine residue on a specific substrate (a biotinylated peptide). Biotinylated peptides that can be used for this purpose include PSK1 (corresponding to amino acids 6-20 of the cell division kinase cdc2-p34) and

PSK2 (corresponding to amino acids 1-17 of gastrin). Both peptides are substrates for a range of tyrosine kinases and are available from Boehringer Mannheim.

The tyrosine kinase reaction is set up by adding the following components in order. First, add 10ul of 5uM Biotinylated Peptide, then 10ul ATP/Mg<sub>2+</sub> (5mM ATP/50mM MgCl<sub>2</sub>), then 10ul of 5x Assay Buffer (40mM imidazole hydrochloride, pH7.3, 40 mM beta-glycerophosphate, 1mM EGTA, 100mM MgCl<sub>2</sub>, 5 mM MnCl<sub>2</sub>, 0.5 mg/ml BSA), then 5ul of Sodium Vanadate(1mM), and then 5ul of water. Mix the components gently and preincubate the reaction mix at 30°C for 2 min. Initiate the reaction by adding 10ul of the control enzyme or the filtered supernatant.

10 The tyrosine kinase assay reaction is then terminated by adding 10 ul of 120mm EDTA and place the reactions on ice.

Tyrosine kinase activity is determined by transferring 50 ul aliquot of reaction mixture to a microtiter plate (MTP) module and incubating at 37°C for 20 min. This allows the streptavidin coated 96 well plate to associate with the biotinylated peptide.

15 Wash the MTP module with 300ul/well of PBS four times. Next add 75 ul of anti-phosphotyrosine antibody conjugated to horse radish peroxidase(anti-P-Tyr-POD(0.5u/ml)) to each well and incubate at 37°C for one hour. Wash the well as above.

Next add 100ul of peroxidase substrate solution (Boehringer Mannheim) and 20 incubate at room temperature for at least 5 mins (up to 30 min). Measure the absorbance of the sample at 405 nm by using ELISA reader. The level of bound peroxidase activity is quantitated using an ELISA reader and reflects the level of tyrosine kinase activity.

25 **Example 20: High-Throughput Screening Assay Identifying Phosphorylation Activity**

As a potential alternative and/or compliment to the assay of protein tyrosine kinase activity described in Example 19, an assay which detects activation (phosphorylation) of major intracellular signal transduction intermediates can also be 30 used. For example, as described below one particular assay can detect tyrosine phosphorylation of the Erk-1 and Erk-2 kinases. However, phosphorylation of other molecules, such as Raf, JNK, p38 MAP, Map kinase kinase (MEK), MEK kinase, Src, Muscle specific kinase (MuSK), IRAK, Tec, and Janus, as well as any other phosphoserine, phosphotyrosine, or phosphothreonine molecule, can be detected by 35 substituting these molecules for Erk-1 or Erk-2 in the following assay.

Specifically, assay plates are made by coating the wells of a 96-well ELISA plate with 0.1ml of protein G (1ug/ml) for 2 hr at room temp, (RT). The plates are then rinsed with PBS and blocked with 3% BSA/PBS for 1 hr at RT. The protein G plates are then treated with 2 commercial monoclonal antibodies (100ng/well) against Erk-1  
5 and Erk-2 (1 hr at RT) (Santa Cruz Biotechnology). (To detect other molecules, this step can easily be modified by substituting a monoclonal antibody detecting any of the above described molecules.) After 3-5 rinses with PBS, the plates are stored at 4°C until use.

A431 cells are seeded at 20,000/well in a 96-well Loprodyn filterplate and  
10 cultured overnight in growth medium. The cells are then starved for 48 hr in basal medium (DMEM) and then treated with EGF (6ng/well) or 50 ul of the supernatants obtained in Example 11 for 5-20 minutes. The cells are then solubilized and extracts filtered directly into the assay plate.

After incubation with the extract for 1 hr at RT, the wells are again rinsed. As a  
15 positive control, a commercial preparation of MAP kinase (10ng/well) is used in place of A431 extract. Plates are then treated with a commercial polyclonal (rabbit) antibody (1ug/ml) which specifically recognizes the phosphorylated epitope of the Erk-1 and Erk-2 kinases (1 hr at RT). This antibody is biotinylated by standard procedures. The bound polyclonal antibody is then quantitated by successive incubations with  
20 Europium-streptavidin and Europium fluorescence enhancing reagent in the Wallac DELFIA instrument (time-resolved fluorescence). An increased fluorescent signal over background indicates a phosphorylation.

**Example 21: Method of Determining Alterations in a Gene  
Corresponding to a Polynucleotide**

RNA isolated from entire families or individual patients presenting with a phenotype of interest (such as a disease) is isolated. cDNA is then generated from these RNA samples using protocols known in the art. (See, Sambrook.) The cDNA is then used as a template for PCR, employing primers surrounding regions of interest in  
30 SEQ ID NO:X. Suggested PCR conditions consist of 35 cycles at 95°C for 30 seconds; 60-120 seconds at 52-58°C; and 60-120 seconds at 70°C, using buffer solutions described in Sidransky, D., et al., Science 252:706 (1991).

PCR products are then sequenced using primers labeled at their 5' end with T4 polynucleotide kinase, employing SequiTHERM Polymerase. (Epicentre Technologies).  
35 The intron-exon borders of selected exons is also determined and genomic PCR

products analyzed to confirm the results. PCR products harboring suspected mutations is then cloned and sequenced to validate the results of the direct sequencing.

PCR products is cloned into T-tailed vectors as described in Holton, T.A. and Graham, M.W., Nucleic Acids Research, 19:1156 (1991) and sequenced with T7 5 polymerase (United States Biochemical). Affected individuals are identified by mutations not present in unaffected individuals.

Genomic rearrangements are also observed as a method of determining alterations in a gene corresponding to a polynucleotide. Genomic clones isolated according to Example 2 are nick-translated with digoxigeninideoxy-uridine 5'- 10 triphosphate (Boehringer Manheim), and FISH performed as described in Johnson, Cg. et al., Methods Cell Biol. 35:73-99 (1991). Hybridization with the labeled probe is carried out using a vast excess of human cot-1 DNA for specific hybridization to the corresponding genomic locus.

Chromosomes are counterstained with 4,6-diamino-2-phenylidole and 15 propidium iodide, producing a combination of C- and R-bands. Aligned images for precise mapping are obtained using a triple-band filter set (Chroma Technology, Brattleboro, VT) in combination with a cooled charge-coupled device camera (Photometrics, Tucson, AZ) and variable excitation wavelength filters. (Johnson, Cv. et al., Genet. Anal. Tech. Appl., 8:75 (1991).) Image collection, analysis and 20 chromosomal fractional length measurements are performed using the ISee Graphical Program System. (Inovision Corporation, Durham, NC.) Chromosome alterations of the genomic region hybridized by the probe are identified as insertions, deletions, and translocations. These alterations are used as a diagnostic marker for an associated disease.

25

**Example 22: Method of Detecting Abnormal Levels of a Polypeptide in a Biological Sample**

A polypeptide of the present invention can be detected in a biological sample, and if an increased or decreased level of the polypeptide is detected, this polypeptide is 30 a marker for a particular phenotype. Methods of detection are numerous, and thus, it is understood that one skilled in the art can modify the following assay to fit their particular needs.

For example, antibody-sandwich ELISAs are used to detect polypeptides in a sample, preferably a biological sample. Wells of a microtiter plate are coated with 35 specific antibodies, at a final concentration of 0.2 to 10 ug/ml. The antibodies are either monoclonal or polyclonal and are produced by the method described in Example 10.

The wells are blocked so that non-specific binding of the polypeptide to the well is reduced.

5 The coated wells are then incubated for > 2 hours at RT with a sample containing the polypeptide. Preferably, serial dilutions of the sample should be used to validate results. The plates are then washed three times with deionized or distilled water to remove unbounded polypeptide.

10 Next, 50 ul of specific antibody-alkaline phosphatase conjugate, at a concentration of 25-400 ng, is added and incubated for 2 hours at room temperature. The plates are again washed three times with deionized or distilled water to remove unbounded conjugate.

15 Add 75 ul of 4-methylumbelliferyl phosphate (MUP) or p-nitrophenyl phosphate (NPP) substrate solution to each well and incubate 1 hour at room temperature. Measure the reaction by a microtiter plate reader. Prepare a standard curve, using serial dilutions of a control sample, and plot polypeptide concentration on the X-axis (log scale) and fluorescence or absorbance of the Y-axis (linear scale).  
Interpolate the concentration of the polypeptide in the sample using the standard curve.

#### Example 23: Formulating a Polypeptide

20 The secreted polypeptide composition will be formulated and dosed in a fashion consistent with good medical practice, taking into account the clinical condition of the individual patient (especially the side effects of treatment with the secreted polypeptide alone), the site of delivery, the method of administration, the scheduling of administration, and other factors known to practitioners. The "effective amount" for purposes herein is thus determined by such considerations.

25 As a general proposition, the total pharmaceutically effective amount of secreted polypeptide administered parenterally per dose will be in the range of about 1 µg/kg/day to 10 mg/kg/day of patient body weight, although, as noted above, this will be subject to therapeutic discretion. More preferably, this dose is at least 0.01 mg/kg/day, and most preferably for humans between about 0.01 and 1 mg/kg/day for the hormone. If 30 given continuously, the secreted polypeptide is typically administered at a dose rate of about 1 µg/kg/hour to about 50 µg/kg/hour, either by 1-4 injections per day or by continuous subcutaneous infusions, for example, using a mini-pump. An intravenous bag solution may also be employed. The length of treatment needed to observe changes and the interval following treatment for responses to occur appears to vary depending 35 on the desired effect.

Pharmaceutical compositions containing the secreted protein of the invention are administered orally, rectally, parenterally, intracistemally, intravaginally,

intraperitoneally, topically (as by powders, ointments, gels, drops or transdermal patch), buccally, or as an oral or nasal spray. "Pharmaceutically acceptable carrier" refers to a non-toxic solid, semisolid or liquid filler, diluent, encapsulating material or formulation auxiliary of any type. The term "parenteral" as used herein refers to modes 5 of administration which include intravenous, intramuscular, intraperitoneal, intrasternal, subcutaneous and intraarticular injection and infusion.

The secreted polypeptide is also suitably administered by sustained-release systems. Suitable examples of sustained-release compositions include semi-permeable polymer matrices in the form of shaped articles, e.g., films, or microcapsules. 10 Sustained-release matrices include polylactides (U.S. Pat. No. 3,773,919, EP 58,481), copolymers of L-glutamic acid and gamma-ethyl-L-glutamate (Sidman, U. et al., Biopolymers 22:547-556 (1983)), poly (2-hydroxyethyl methacrylate) (R. Langer et al., J. Biomed. Mater. Res. 15:167-277 (1981), and R. Langer, Chem. Tech. 12:98-105 (1982)), ethylene vinyl acetate (R. Langer et al.) or poly-D-(-)-3-hydroxybutyric 15 acid (EP 133,988). Sustained-release compositions also include liposomally entrapped polypeptides. Liposomes containing the secreted polypeptide are prepared by methods known per se: DE 3,218,121; Epstein et al., Proc. Natl. Acad. Sci. USA 82:3688-3692 (1985); Hwang et al., Proc. Natl. Acad. Sci. USA 77:4030-4034 (1980); EP 52,322; EP 36,676; EP 88,046; EP 143,949; EP 142,641; Japanese Pat. Appl. 83-118008; 20 U.S. Pat. Nos. 4,485,045 and 4,544,545; and EP 102,324. Ordinarily, the liposomes are of the small (about 200-800 Angstroms) unilamellar type in which the lipid content is greater than about 30 mol. percent cholesterol, the selected proportion being adjusted for the optimal secreted polypeptide therapy.

For parenteral administration, in one embodiment, the secreted polypeptide is 25 formulated generally by mixing it at the desired degree of purity, in a unit dosage injectable form (solution, suspension, or emulsion), with a pharmaceutically acceptable carrier, i.e., one that is non-toxic to recipients at the dosages and concentrations employed and is compatible with other ingredients of the formulation. For example, the formulation preferably does not include oxidizing agents and other compounds that are 30 known to be deleterious to polypeptides.

Generally, the formulations are prepared by contacting the polypeptide uniformly and intimately with liquid carriers or finely divided solid carriers or both. Then, if necessary, the product is shaped into the desired formulation. Preferably the carrier is a parenteral carrier, more preferably a solution that is isotonic with the blood 35 of the recipient. Examples of such carrier vehicles include water, saline, Ringer's solution, and dextrose solution. Non-aqueous vehicles such as fixed oils and ethyl oleate are also useful herein, as well as liposomes.

The carrier suitably contains minor amounts of additives such as substances that enhance isotonicity and chemical stability. Such materials are non-toxic to recipients at the dosages and concentrations employed, and include buffers such as phosphate, citrate, succinate, acetic acid, and other organic acids or their salts; antioxidants such as 5 ascorbic acid; low molecular weight (less than about ten residues) polypeptides, e.g., polyarginine or tripeptides; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids, such as glycine, glutamic acid, aspartic acid, or arginine; monosaccharides, 10 disaccharides, and other carbohydrates including cellulose or its derivatives, glucose, manose, or dextrans; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; counterions such as sodium; and/or nonionic surfactants such as polysorbates, poloxamers, or PEG.

15 The secreted polypeptide is typically formulated in such vehicles at a concentration of about 0.1 mg/ml to 100 mg/ml, preferably 1-10 mg/ml, at a pH of about 3 to 8. It will be understood that the use of certain of the foregoing excipients, carriers, or stabilizers will result in the formation of polypeptide salts.

20 Any polypeptide to be used for therapeutic administration can be sterile. Sterility is readily accomplished by filtration through sterile filtration membranes (e.g., 0.2 micron membranes). Therapeutic polypeptide compositions generally are placed into a container having a sterile access port, for example, an intravenous solution bag or vial having a stopper pierceable by a hypodermic injection needle.

25 Polypeptides ordinarily will be stored in unit or multi-dose containers, for example, sealed ampoules or vials, as an aqueous solution or as a lyophilized formulation for reconstitution. As an example of a lyophilized formulation, 10-ml vials are filled with 5 ml of sterile-filtered 1% (w/v) aqueous polypeptide solution, and the resulting mixture is lyophilized. The infusion solution is prepared by reconstituting the lyophilized polypeptide using bacteriostatic Water-for-Injection.

30 The invention also provides a pharmaceutical pack or kit comprising one or more containers filled with one or more of the ingredients of the pharmaceutical compositions of the invention. Associated with such container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, which notice reflects approval by the agency of manufacture, use or sale for human administration. In addition, the polypeptides of the present invention may be employed in conjunction with other therapeutic compounds.

It will be appreciated that conditions caused by a decrease in the standard or normal expression level of a secreted protein in an individual can be treated by administering the polypeptide of the present invention, preferably in the secreted form. Thus, the invention also provides a method of treatment of an individual in need of an increased level of the polypeptide comprising administering to such an individual a pharmaceutical composition comprising an amount of the polypeptide to increase the activity level of the polypeptide in such an individual.

For example, a patient with decreased levels of a polypeptide receives a daily dose 0.1-100 ug/kg of the polypeptide for six consecutive days. Preferably, the polypeptide is in the secreted form. The exact details of the dosing scheme, based on administration and formulation, are provided in Example 23.

#### Example 25: Method of Treating Increased Levels of the Polypeptide

Antisense technology is used to inhibit production of a polypeptide of the present invention. This technology is one example of a method of decreasing levels of a polypeptide, preferably a secreted form, due to a variety of etiologies, such as cancer.

For example, a patient diagnosed with abnormally increased levels of a polypeptide is administered intravenously antisense polynucleotides at 0.5, 1.0, 1.5, 2.0 and 3.0 mg/kg day for 21 days. This treatment is repeated after a 7-day rest period if the treatment was well tolerated. The formulation of the antisense polynucleotide is provided in Example 23.

#### Example 26: Method of Treatment Using Gene Therapy

One method of gene therapy transplants fibroblasts, which are capable of expressing a polypeptide, onto a patient. Generally, fibroblasts are obtained from a subject by skin biopsy. The resulting tissue is placed in tissue-culture medium and separated into small pieces. Small chunks of the tissue are placed on a wet surface of a tissue culture flask, approximately ten pieces are placed in each flask. The flask is turned upside down, closed tight and left at room temperature over night. After 24 hours at room temperature, the flask is inverted and the chunks of tissue remain fixed to the bottom of the flask and fresh media (e.g., Ham's F12 media, with 10% FBS, penicillin and streptomycin) is added. The flasks are then incubated at 37°C for approximately one week.

At this time, fresh media is added and subsequently changed every several days. After an additional two weeks in culture, a monolayer of fibroblasts emerge. The monolayer is trypsinized and scaled into larger flasks.

pMV-7 (Kirschmeier, P.T. et al., DNA, 7:219-25 (1988)), flanked by the long terminal repeats of the Moloney murine sarcoma virus, is digested with EcoRI and HindIII and subsequently treated with calf intestinal phosphatase. The linear vector is fractionated on agarose gel and purified, using glass beads.

5       The cDNA encoding a polypeptide of the present invention can be amplified using PCR primers which correspond to the 5' and 3' end sequences respectively as set forth in Example 1. Preferably, the 5' primer contains an EcoRI site and the 3' primer includes a HindIII site. Equal quantities of the Moloney murine sarcoma virus linear backbone and the amplified EcoRI and HindIII fragment are added together, in the  
10 presence of T4 DNA ligase. The resulting mixture is maintained under conditions appropriate for ligation of the two fragments. The ligation mixture is then used to transform bacteria HB101, which are then plated onto agar containing kanamycin for the purpose of confirming that the vector has the gene of interest properly inserted.

15      The amphotropic pA317 or GP+am12 packaging cells are grown in tissue culture to confluent density in Dulbecco's Modified Eagles Medium (DMEM) with 10% calf serum (CS), penicillin and streptomycin. The MSV vector containing the gene is then added to the media and the packaging cells transduced with the vector. The packaging cells now produce infectious viral particles containing the gene (the packaging cells are now referred to as producer cells).

20      Fresh media is added to the transduced producer cells, and subsequently, the media is harvested from a 10 cm plate of confluent producer cells. The spent media, containing the infectious viral particles, is filtered through a millipore filter to remove detached producer cells and this media is then used to infect fibroblast cells. Media is removed from a sub-confluent plate of fibroblasts and quickly replaced with the media  
25 from the producer cells. This media is removed and replaced with fresh media. If the titer of virus is high, then virtually all fibroblasts will be infected and no selection is required. If the titer is very low, then it is necessary to use a retroviral vector that has a selectable marker, such as neo or his. Once the fibroblasts have been efficiently infected, the fibroblasts are analyzed to determine whether protein is produced.

30      The engineered fibroblasts are then transplanted onto the host, either alone or after having been grown to confluence on cytodex 3 microcarrier beads.

**Example 27: Method of Treatment Using Gene Therapy - In Vivo**

35      Another aspect of the present invention is using *in vivo* gene therapy methods to treat disorders, diseases and conditions. The gene therapy method relates to the introduction of naked nucleic acid (DNA, RNA, and antisense DNA or RNA) sequences into an animal to increase or decrease the expression

of the polypeptide of the present invention. A polynucleotide of the present invention may be operatively linked to a promoter or any other genetic elements necessary for the expression of the encoded polypeptide by the target tissue. Such gene therapy and delivery techniques and methods are known in the art, 5 see, for example, WO90/11092, WO98/11779; U.S. Patent NO. 5693622, 5705151, 5580859; Tabata H. et al. (1997) *Cardiovasc. Res.* 35(3):470-479, Chao J et al. (1997) *Pharmacol. Res.* 35(6):517-522, Wolff J.A. (1997) *Neuromuscul. Disord.* 7(5):314-318, Schwartz B. et al. (1996) *Gene Ther.* 3(5):405-411, Tsurumi Y. et al. (1996) *Circulation* 94(12):3281-3290. 10 (incorporated herein by reference).

The polynucleotide constructs of the present invention may be delivered by any method that delivers injectable materials to the cells of an animal, such as, injection into the interstitial space of tissues (heart, muscle, skin, lung, liver, intestine and the like). These polynucleotide constructs can be delivered in a 15 pharmaceutically acceptable liquid or aqueous carrier.

The term "naked" polynucleotide, DNA or RNA, refers to sequences that are free from any delivery vehicle that acts to assist, promote, or facilitate entry into the cell, including viral sequences, viral particles, liposome formulations, lipofectin or precipitating agents and the like. However, the 20 polynucleotides may also be delivered in liposome formulations (such as those taught in Felgner P.L. et al. (1995) *Ann. NY Acad. Sci.* 772:126-139 and Abdallah B. et al. (1995) *Biol. Cell* 85(1):1-7) which can be prepared by methods well known to those skilled in the art.

The polynucleotide vector constructs of the present invention used in 25 the gene therapy method are preferably constructs that will not integrate into the host genome nor will they contain sequences that allow for replication. Any strong promoter known to those skilled in the art can be used for driving the expression of DNA. Unlike other gene therapies techniques, one major advantage of introducing naked nucleic acid sequences into target cells is the 30 transitory nature of the polynucleotide synthesis in the cells. Studies have shown that non-replicating DNA sequences can be introduced into cells to provide production of the desired polypeptide for periods of up to six months.

The polynucleotide construct of the present invention can be delivered to 35 the interstitial space of tissues within the an animal, including of muscle, skin, brain, lung, liver, spleen, bone marrow, thymus, heart, lymph, blood, bone, cartilage, pancreas, kidney, gall bladder, stomach, intestine, testis, ovary, uterus, rectum, nervous system, eye, gland, and connective tissue. Interstitial

space of the tissues comprises the intercellular fluid, mucopolysaccharide matrix among the reticular fibers of organ tissues, elastic fibers in the walls of vessels or chambers, collagen fibers of fibrous tissues, or that same matrix within connective tissue ensheathing muscle cells or in the lacunae of bone. It is similarly the space occupied by the plasma of the circulation and the lymph fluid of the lymphatic channels. Delivery to the interstitial space of muscle tissue is preferred for the reasons discussed below. They may be conveniently delivered by injection into the tissues comprising these cells. They are preferably delivered to and expressed in persistent, non-dividing cells which are differentiated, although delivery and expression may be achieved in non-differentiated or less completely differentiated cells, such as, for example, stem cells of blood or skin fibroblasts. *In vivo* muscle cells are particularly competent in their ability to take up and express polynucleotides.

For the naked polynucleotide injection, an effective dosage amount of DNA or RNA will be in the range of from about 0.05 g/kg body weight to about 50 mg/kg body weight. Preferably the dosage will be from about 0.005 mg/kg to about 20 mg/kg and more preferably from about 0.05 mg/kg to about 5 mg/kg. Of course, as the artisan of ordinary skill will appreciate, this dosage will vary according to the tissue site of injection. The appropriate and effective dosage of nucleic acid sequence can readily be determined by those of ordinary skill in the art and may depend on the condition being treated and the route of administration. The preferred route of administration is by the parenteral route of injection into the interstitial space of tissues. However, other parenteral routes may also be used, such as, inhalation of an aerosol formulation particularly for delivery to lungs or bronchial tissues, throat or mucous membranes of the nose. In addition, naked polynucleotide constructs can be delivered to arteries during angioplasty by the catheter used in the procedure.

The dose response effects of injected polynucleotide in muscle *in vivo* is determined as follows. Suitable template DNA for production of mRNA coding for the polypeptide of the present invention is prepared in accordance with a standard recombinant DNA methodology. The template DNA, which may be either circular or linear, is either used as naked DNA or complexed with liposomes. The quadriceps muscles of mice are then injected with various amounts of the template DNA.

Five to six week old female and male Balb/C mice are anesthetized by intraperitoneal injection with 0.3 ml of 2.5% Avertin. A 1.5 cm incision is made on the anterior thigh, and the quadriceps muscle is directly visualized. The

template DNA is injected in 0.1 ml of carrier in a 1 cc syringe through a 27 gauge needle over one minute, approximately 0.5 cm from the distal insertion site of the muscle into the knee and about 0.2 cm deep. A suture is placed over the injection site for future localization, and the skin is closed with stainless steel

5 clips.

After an appropriate incubation time (e.g., 7 days) muscle extracts are prepared by excising the entire quadriceps. Every fifth 15 um cross-section of the individual quadriceps muscles is histochemically stained for protein expression. A time course for protein expression may be done in a similar fashion except that quadriceps from  
10 different mice are harvested at different times. Persistence of DNA in muscle following injection may be determined by Southern blot analysis after preparing total cellular DNA and HIRT supernatants from injected and control mice. The results of the above experimentation in mice can be used to extrapolate proper dosages and other treatment parameters in humans and other animals using naked DNA of the present invention.

15 It will be clear that the invention may be practiced otherwise than as particularly described in the foregoing description and examples. Numerous modifications and variations of the present invention are possible in light of the above teachings and, therefore, are within the scope of the appended claims.

20 The entire disclosure of each document cited (including patents, patent applications, journal articles, abstracts, laboratory manuals, books, or other disclosures) in the Background of the Invention, Detailed Description, and Examples is hereby incorporated herein by reference.

Further, the Sequence Listing submitted herewith in paper and computer readable forms are herein incorporated by reference in their entireties.

## INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description  
on page 79, line N/A

## B. IDENTIFICATION OF DEPOSIT

Further deposits are identified on an additional sheet 

Name of depositary institution  
American Type Culture Collection

Address of depositary institution (*including postal code and country*)

10801 University Boulevard  
Manassas, Virginia 20110-2209  
United States of America

Date of deposit July 3, 1997Accession Number 209138

C. ADDITIONAL INDICATIONS (*leave blank if not applicable*) This information is continued on an additional sheet

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (*if the indications are not for all designated States*)E. SEPARATE FURNISHING OF INDICATIONS (*leave blank if not applicable*)

The indications listed below will be submitted to the International Bureau later (*specify the general nature of the indications, e.g., "Accession Number of Deposit"*)

For receiving Office use only

This sheet was received with the international application

For International Bureau use only

This sheet was received by the International Bureau on:

Authorized officer



Authorized officer

## INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

<b>A. The indications made below relate to the microorganism referred to in the description on page <u>80</u>, line <u>N/A</u></b>		
<b>B. IDENTIFICATION OF DEPOSIT</b>		Further deposits are identified on an additional sheet <input type="checkbox"/>
Name of depositary institution American Type Culture Collection		
Address of depositary institution ( <i>including postal code and country</i> ) 10801 University Boulevard Manassas, Virginia 20110-2209 United States of America		
Date of deposit	July 3, 1997	Accession Number 209139
<b>C. ADDITIONAL INDICATIONS</b> ( <i>leave blank if not applicable</i> )		This information is continued on an additional sheet <input type="checkbox"/>
<b>D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE</b> ( <i>if the indications are not for all designated States</i> )		
<b>E. SEPARATE FURNISHING OF INDICATIONS</b> ( <i>leave blank if not applicable</i> ) The indications listed below will be submitted to the International Bureau later ( <i>specify the general nature of the indications, e.g., "Accession Number of Deposit"</i> )		

For receiving Office use only



This sheet was received with the international application

For International Bureau use only



This sheet was received by the International Bureau on:

Authorized officer

Virginia L. Lily

Authorized officer

## INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description  
on page 82, line N/A

## B. IDENTIFICATION OF DEPOSIT

Further deposits are identified on an additional sheet

Name of depositary institution:

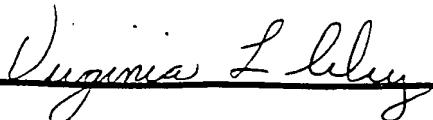
American Type Culture Collection

Address of depositary institution (*including postal code and country*)10801 University Boulevard  
Manassas, Virginia 20110-2209  
United States of AmericaDate of deposit July 9, 1997Accession Number 209141C. ADDITIONAL INDICATIONS (*leave blank if not applicable*) This information is continued on an additional sheetD. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (*if the indications are not for all designated States*)E. SEPARATE FURNISHING OF INDICATIONS (*leave blank if not applicable*)The indications listed below will be submitted to the International Bureau later (*specify the general nature of the indications, e.g., "Accession Number of Deposit"*)

For receiving Office use only

This sheet was received with the international application

Authorized officer



For International Bureau use only

This sheet was received by the International Bureau on:

Authorized officer

***What Is Claimed Is:***

1. An isolated nucleic acid molecule comprising a polynucleotide having a nucleotide sequence at least 95% identical to a sequence selected from the group consisting of:
  - (a) a polynucleotide fragment of SEQ ID NO:X or a polynucleotide fragment of the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
  - (b) a polynucleotide encoding a polypeptide fragment of SEQ ID NO:Y or a polypeptide fragment encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
  - (c) a polynucleotide encoding a polypeptide domain of SEQ ID NO:Y or a polypeptide domain encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
  - (d) a polynucleotide encoding a polypeptide epitope of SEQ ID NO:Y or a polypeptide epitope encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
  - (e) a polynucleotide encoding a polypeptide of SEQ ID NO:Y or the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X, having biological activity;
  - (f) a polynucleotide which is a variant of SEQ ID NO:X;
  - (g) a polynucleotide which is an allelic variant of SEQ ID NO:X;
  - (h) a polynucleotide which encodes a species homologue of the SEQ ID NO:Y;
  - (i) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(h), wherein said polynucleotide does not hybridize under stringent conditions to a nucleic acid molecule having a nucleotide sequence of only A residues or of only T residues.
2. The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment comprises a nucleotide sequence encoding a secreted protein.
3. The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment comprises a nucleotide sequence encoding the sequence identified as SEQ ID NO:Y or the polypeptide encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X.

4. The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment comprises the entire nucleotide sequence of SEQ ID NO:X or the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X.

5

5. The isolated nucleic acid molecule of claim 2, wherein the nucleotide sequence comprises sequential nucleotide deletions from either the C-terminus or the N-terminus.

10

6. The isolated nucleic acid molecule of claim 3, wherein the nucleotide sequence comprises sequential nucleotide deletions from either the C-terminus or the N-terminus.

15

7. A recombinant vector comprising the isolated nucleic acid molecule of claim 1.

8. A method of making a recombinant host cell comprising the isolated nucleic acid molecule of claim 1.

20

9. A recombinant host cell produced by the method of claim 8.

10. The recombinant host cell of claim 9 comprising vector sequences.

11. An isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence selected from the group consisting of:

(a) a polypeptide fragment of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;

(b) a polypeptide fragment of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z, having biological activity;

30 (c) a polypeptide domain of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;

(d) a polypeptide epitope of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;

(e) a secreted form of SEQ ID NO:Y or the encoded sequence included in

35 ATCC Deposit No:Z;

(f) a full length protein of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;

- (g) a variant of SEQ ID NO:Y;
- (h) an allelic variant of SEQ ID NO:Y; or
- (i) a species homologue of the SEQ ID NO:Y.

12. The isolated polypeptide of claim 11, wherein the secreted form or the  
5 full length protein comprises sequential amino acid deletions from either the C-terminus  
or the N-terminus.

13. An isolated antibody that binds specifically to the isolated polypeptide of  
claim 11.

10 14. A recombinant host cell that expresses the isolated polypeptide of claim  
11.

15 15. A method of making an isolated polypeptide comprising:  
(a) culturing the recombinant host cell of claim 14 under conditions such that  
said polypeptide is expressed; and  
(b) recovering said polypeptide.

20 16. The polypeptide produced by claim 15.

17. A method for preventing, treating, or ameliorating a medical condition,  
comprising administering to a mammalian subject a therapeutically effective amount of  
the polypeptide of claim 11 or the polynucleotide of claim 1.

25 18. A method of diagnosing a pathological condition or a susceptibility to a  
pathological condition in a subject comprising:  
(a) determining the presence or absence of a mutation in the polynucleotide of  
claim 1; and  
(b) diagnosing a pathological condition or a susceptibility to a pathological  
30 condition based on the presence or absence of said mutation.

19. A method of diagnosing a pathological condition or a susceptibility to a  
pathological condition in a subject comprising:  
(a) determining the presence or amount of expression of the polypeptide of  
35 claim 11 in a biological sample; and  
(b) diagnosing a pathological condition or a susceptibility to a pathological  
condition based on the presence or amount of expression of the polypeptide.

20. A method for identifying a binding partner to the polypeptide of claim 11 comprising:

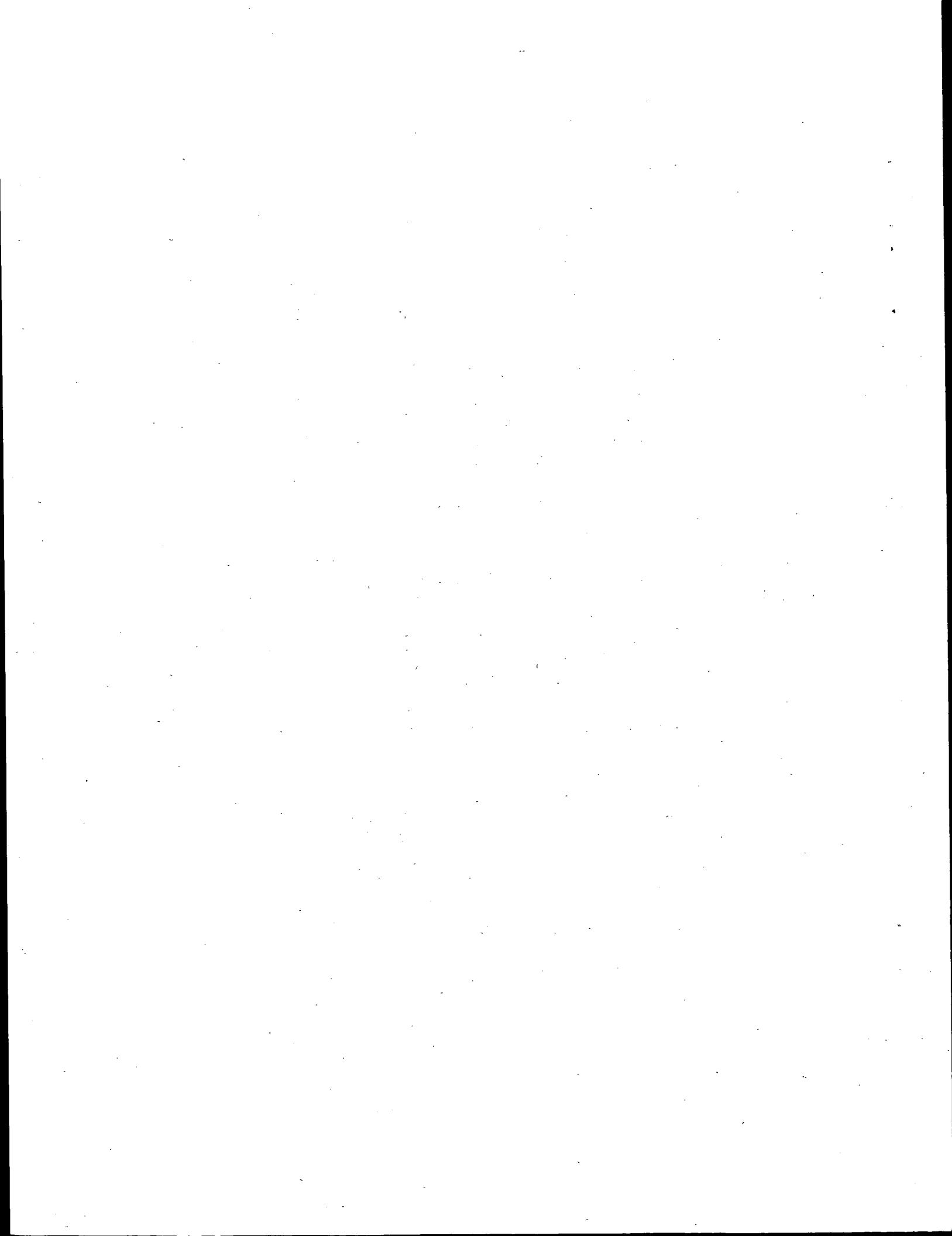
- (a) contacting the polypeptide of claim 11 with a binding partner; and
- 5 (b) determining whether the binding partner effects an activity of the polypeptide.

21. The gene corresponding to the cDNA sequence of SEQ ID NO:Y.

10 22. A method of identifying an activity in a biological assay, wherein the method comprises:

- (a) expressing SEQ ID NO:X in a cell;
- (b) isolating the supernatant;
- (c) detecting an activity in a biological assay; and
- 15 (d) identifying the protein in the supernatant having the activity.

23. The product produced by the method of claim 22.



<110> Human Genome Sciences, Inc.  
<120> 64 Human Secreted Proteins  
5 <130> PZ011PCT  
  
<140> Unassigned  
<141> 1998-07-15  
10 <150> 60/052,870  
<151> 1997-07-16  
  
<150> 60/052,871  
15 <151> 1997-07-16  
  
<150> 60/052,872  
<151> 1997-07-16  
  
20 <150> 60/052,661  
<151> 1997-07-16  
  
<150> 60/052,874  
<151> 1997-07-16  
25 <150> 60/052,873  
<151> 1997-07-16  
  
<150> 60/052,875  
30 <151> 1997-07-16  
  
<150> 60/053,440  
<151> 1997-07-22  
  
35 <150> 60/053,441  
<151> 1997-07-22  
  
<150> 60/053,442  
<151> 1997-07-22  
40 <150> 60/055,952  
<151> 1997-08-18  
  
<150> 60/055,725  
45 <151> 1997-08-18  
  
<150> 60/056,359  
<151> 1997-08-18  
  
50 <150> 60/055,985  
<151> 1997-08-18  
  
<150> 60/055,724  
<151> 1997-08-18  
55

<150> 60/055,726  
 <151> 1997-08-18

5 <150> 60/056,361  
 <151> 1997-08-18

<150> 60/055,989  
 <151> 1997-08-18

10 <150> 60/055,946  
 <151> 1997-08-18

<150> 60/055,683  
 <151> 1997-08-18

15 <160> 178  
 <170> PatentIn Ver. 2.0

20 <210> 1  
 <211> 733  
 <212> DNA  
 <213> Homo sapiens

25 <400> 1  
 gggatccgga gcccaaattct tctgacaaaa ctcacacatg cccaccgtgc ccagcacctg 60  
 aattcggagg tgacccgtca gtcttccctct tccccccaaa acccaaggac accctcatga 120  
 tctcccgAAC tcctgaggTC acatgcgtgg tggtggacgt aagccacgaa gaccctgagg 180  
 tcaagttcaa ctggtaCGTg gacggcgtgg aggtgcataa tgccaagaca aagcccgccc 240  
 aggagcagta caacacgacg taccgtgtgg tcagcgtctt caccgtcctg caccaggact 300  
 ggctgaatgg caaggagtac aagtgcacgg tctccaacaa agccctccca acccccatcg 360  
 agaaaaccat ctccaaAGCC aaagggcagc cccgagaacc acaggtgtac accctgcccc 420  
 catcccgggA tgagctgacc aagaaccagg tcagcctgac ctgcctggTC aaaggcttct 480  
 30 atccaagcga catcggcgtg gagtgggaga gcaatggca gccggagaac aactacaaga 540  
 ccacgcctcc cgtgcgtggAC tccgacggct ccttcttctt ctacagcaag ctcaccgtgg 600  
 acaagagcag gtggcagcag gggAACgttct tctcatgctc cgtgatgcat gaggctctgc 660  
 acaaccacta cacgcagaag agcctctccc tgcctccggg taaatgagtg cgacggccgc 720  
 gactcttagag gat 733

40 <210> 2  
 <211> 5  
 <212> PRT  
 45 <213> Homo sapiens

<220>  
 <221> Site  
 <222> (3)

50 <223> Xaa equals any one of the twenty naturally occurring L-amino acids

<400> 2  
 Trp Ser Xaa Trp Ser  
 1 5

55 <210> 3  
 <211> 86  
 <212> DNA

<213> Homo sapiens

5 <400> 3  
gcccctcgag attccccga aatcttagatt tcccccgaat gattccccg aaatgattc  
cccgaataat ctggcatctc aattag 60  
86

10 <210> 4  
<211> 27  
<212> DNA  
<213> Homo sapiens

15 <400> 4  
gcggcaagct ttttgc当地 cctaggc 27

20 <210> 5  
<211> 271  
<212> DNA  
<213> Homo sapiens

25 <400> 5  
ctcgagattt ccccgaaatc tagatttccc cgaaatgatt tcccccgaat gattccccg  
aaatatctgc catctcaatt agtcagcaac catagccccg cccctaactc cgcccatccc 60  
gcccctaact cc当地 cccatc tccgccccat ggctgactaa ttttttttat 120  
ttatgc当地 gcccaggccg cctcgccctc tgagctattc cagaagtatg gaggaggctt 180  
tttggaggc ctaggctttt gcaaaaagct t 240  
271

30 <210> 6  
<211> 32  
<212> DNA  
<213> Homo sapiens

35 <400> 6  
gc当地 cggagg gatgacagcg atagaacccc gg 32

40 <210> 7  
<211> 31  
<212> DNA  
<213> Homo sapiens

45 <400> 7  
gc当地 agcttc gc当地 actcccc ggatccgcct c 31

50 <210> 8  
<211> 12  
<212> DNA  
<213> Homo sapiens

55 <400> 8  
ggggactttc cc 12

60 <210> 9  
<211> 73  
<212> DNA  
<213> Homo sapiens

	<400> 9	
	gcggccctcg a gggactttc ccggggactt tccgggact ttccgggact ttccatcctg	60
	ccatctcaat tag	73
5		
	<210> 10	
	<211> 256	
	<212> DNA	
10	<213> Homo sapiens	
	<400> 10	
	ctcgagggg a cttccggg gactttccg ggactttcca tctgccccatct	60
	caatttagtca gcaaccatag tcccgcccc aactccgccc atcccgcccc taactccgccc	120
15	cagttccggc cattctccgc cccatggctg actaatttt tttattttatg cagaggccga	180
	ggccgcctcg gcctctgagc tattccagaa gtatgtgagga ggctttttt gaggcctagg	240
	cttttgcaaa aagctt	256
20	<210> 11	
	<211> 558	
	<212> DNA	
	<213> Homo sapiens	
25	<220>	
	<221> SITE	
	<222> (546)	
	<223> n equals a,t,g, or c	
30	<400> 11	
	gaattcggca cgagctggc tgca gttggc gattccgc ggtgaaagca gccagtgc	60
	agggtctttt cctgagtgca cctggccctg ccgcggcgc atgcccattgg gtcgtgcgc	120
	gcttttctac ttgcgcgc ctca ctgc ggtgtactgg gagggtaacc tggaggcggt	180
35	gccttttattt ttccgaaccgc ccgc tca ctgc agacagtggc tagaagtgtc tcttggac	240
	gtgagtttagc ctta acctgt tatgccccca gagccctcag tggagcgc gtactttgcc	300
	ggcatgcagt ttgattttcc ggtgataatc cgacgagttt gacagattga ggtatgtgac	360
	aaagttggcc gtcagttggc ggcacttga ctgcgtgcgg accctggcc tgcctttggaa	420
	agagatagtg ttcttagggc tggtttact gtccttaag actgaarggt ggarctggga	480
40	tatagatgtg ttgtttctt tcaa atcaaa cctgc ttttgc tgcactcg aaggggggccc	540
	cggtanccaa ttgc	558
	<210> 12	
	<211> 715	
45	<212> DNA	
	<213> Homo sapiens	
	<400> 12	
	cggatttcga gtgctttctt ctttacctcc accctcccccc atgttttaat gcagccctcc	60
50	aaaaatattt actgagtgta gactcttagac cagggccctgt gctaggataac aaagatgaat	120
	gaggcaccac ctttatcttc gaggtagata ttttttattt tttttccctgt	180
	ctgcctccctt tgagtagtac atgttttagt aaggggaaca gacactaaag agtccctggta	240
	atgatgagca aagtactgca tgagtaagta tctggggggc aagtgtcccc actaggactc	300
55	ctgtcagatc tggaaaaggc ctgaggaatc tgatacatga cttaatgcag cgtataacttg	360
	cagcctggaa aactaaatggaa tgacaaaata gacattcttgc tca gttgtgag ccattctctg	420
	agtccmaggg gactacataa ttcaaaaccag aattggcat tttggagttt gcactcttag	480
	cagtatacag tggatgtaaa tttaagaatc aattttat tttttcagtt tttatgtaca	540
	taaaaacctgc ttactacaag agacccagtt tatttttttgc ttttttttgc aatcataaag	600
60	tatatttcat cataataagg ctccgtaaa ttgttcat ttttttttgc caatcaaagac	660
	atatatctga aaataaaatgt tcctgaacctt gaaaaaaaaaaa aaaaaaaaaaaa ctcga	715

5           <210> 13  
 <211> 838  
 <212> DNA  
 <213> Homo sapiens

10          <400> 13  
 gaattcggca cgagccaaaa caaaagaaaac ctttggaggc atgtgtcaga acagagaaaag         60  
 tgcctgggtt ttgcttatag aatcaaataat gttctcattc tacctactgt tttcattcta         120  
 catagtgttt tccttcttta tagttttacg tccttccttcc aggaatgagt ctattaagaa         180  
 aataggtgtt atcttttagc ttggcattt gactttcagg ataataagagc tatctgtcac         240  
 tgacagaaaa gctttgacaa gtgttaata ctctggatt accttcattt tacttttgca         300  
 atcattatgt gaacattgtc ttccgtccac atctayaggc tagtawgtaa caccgttgac         360  
 15          taaatccaaa ctttaggcta gggaaaaagg gtatacttc tgggtttcggtt ttgttagatta         420  
 tggtagatc taaycaaaaac aggacagtgg tccaaacaga aaattgttat tttctgtatc         480  
 ttgttaatct agatttgag ttttaagat gaattttatgg ttccctttctt gatatcattt         540  
 ctcatctgca gctcctaataatg cctggtaacctt tgggtatgga gtgaggagag acaatggaca         600  
 gttttatata agaaatggaa gtaatgatac tatcttcctt ggaatatttg caggccccag         660  
 20          aggagatgtat gagcaaggac tggggcctt tattacacac aacagggttg tagttactat         720  
 cccagcaagg aaagggtgtt tctttcttctt ttcatgcaaa ttatctatga tgacctaaca         780  
 gttttagat agtgagtggc ctaaccacaa caataaaaaaa aaaaaaaaaa aactcgaa         838

25          <210> 14  
 <211> 513  
 <212> DNA  
 <213> Homo sapiens

30          <400> 14  
 ctgcaggaaat tcggcacgag ggaacaactc catgtttttt taaaggccta gagaacatat         60  
 atccagtgcc tttcctttt gcctttgtat tcattttt ggaaatttac tggaaagatga         120  
 cggttctggc caaaaggctg gttttgttt tgggtcacat ttcttgctt ctctgcgtt         180  
 35          gaatcttggaa tttagatgtat gacatggta agatctcagc aaccttcattt actagaagat         240  
 catgtggatt ggaatcatac aatggggaaac aatggaaaaa gagtactttt gaaatagtgc         300  
 tggagaccac tggaccaca gaatgtcaag acacgtgctg ccattactgt tactatttgg         360  
 aaaaatacatt ctgttaatg caaccttagg gggtttggagg gggaaagtctg ttgggaaatg         420  
 aattgcaaga aaaatattac accctgaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa         480  
 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaact cga         513

40

45          <210> 15  
 <211> 712  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> SITE  
 <222> (565)

50          <223> n equals a,t,g, or c

<400> 15  
 gaattcggca cgagctggac aggaccggag aggaccccgca gtaaccgcgg aacagacact         60  
 cccggcagcg gcccggcccg cggactgtc acggacgag cccggacgct tggccatggc         120  
 55          ggcccgatcc gcactggcgc tgctgtct gctgccagtc ctgttcctgc cgggtcagag         180  
 ycgctcagag cccgagacca cccggccac ccctacccca atccgggtg gcaactcgtc         240  
 aktgagcagg cccctgcccc gcacgtggct ccacgcctgc gggccatacc ccaaaccagg         300  
 cctgctcatc ctgtggcccc cgctggccct gtggccatt ctctgttagg gacgccccagc         360  
 cagccacctc taatcgccg ctgggactgg cctggccat tgagcaacag agacgcttga         420  
 60          cagccgcccc cctccatcc ttgacttcac ccagaaaatgg gtccagaaaa ctgaatccca         480

	ccagcaactgg	tttggaggcaa	ccggacacccg	aggtttcacc	tccagggrtt	ccatggaaga	540
	gcctcaatgg	agatgccaca	tcctnactga	gttaaagatg	ggctgagggaa	cttgggttacc	600
	cacaagtytg	ccttgggrat	aaaaagaaaa	tattnacctt	tagtttggtt	cattaaatgc	660
	atgaagtcaa	aatatgaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaactc ga	712
5							
	<210> 16						
	<211> 652						
	<212> DNA						
10	<213> Homo sapiens						
	<400> 16						
	gaattcggca	cgagcaacag	tggggcaactc	tgctcccagg	cagggtccccac	tggggctgagc	60
	cgcacagccct	ggctttggc	ttccctgact	gcaccaccca	catcasctgc	ctctagccct	120
15	taamatacaa	aacttcccc	agtcaactggc	cggcaggctg	agttggggga	tgtgttacat	180
	ccctgggtcc	actggggggc	agtgttggcc	atggtgttgg	tgctggctt	gcccagagggc	240
	gttggagttgg	ctgtgtgggg	cggtgagcgc	cggcccagcc	tgatggAAC	cactgtacca	300
	ggcccaggcc	tcagcctctg	agaaggactt	ccctgtgtca	ctcactcata	catgtcctca	360
20	ggacgtgaag	acatttcagc	agacaaaagt	tttccttcgaa	tttccttcga	atcgtccaga	420
	tactttggaga	catctccccc	tcacccgttgg	ggtgctgggg	cagtccctagg	cgtggggggc	480
	gatgggtgg	cagctgtgc	tgccctgctg	gggggtggca	gcccctggag	cacacagtgg	540
	tgaagacatt	cctgaatatg	tctcaggctg	tagaaatctt	attttgttgg	aagatttttag	600
	agaatcatca	aaataaactt	ttaccaaata	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaactc ga	652
25							
	<210> 17						
	<211> 742						
	<212> DNA						
30	<213> Homo sapiens						
	<400> 17						
	ggtcgaccca	cgcgtctgat	atgatctctt	tatccttctc	cctttgattt	atctttttct	60
	tttagctgtat	tttagcttttc	ttcttttctc	tgttagttggc	ggaatcagct	cagttacatt	120
35	ttttactaag	ttacccacat	tctgacactc	cttgacagtk	ttaagatctt	cttctaaacac	180
	acttgaatag	aatggatact	ggaatctatt	ttgacagctg	ttgaaaatct	attctgttgt	240
	tacaggaggt	taaggaggtt	atttgttaca	ctgggatttat	ttaatgttacc	ttttgaaaag	300
	gtgtgcagac	tgttcaggca	aatagtattt	tttagaatta	aatgattttt	gttttcacag	360
	ttaaatttac	aatatgtatg	cttttaagaa	ttatacacct	agtaatattt	ttcattaatt	420
40	tctccaccag	tgttagtaata	gtacattaca	atgttctcaa	ttaccggc	cttctaaaat	480
	gcaggtgttag	agtcytaaaa	tacagctagt	ctatkgccag	ctgtccata	gataacccctc	540
	tcyttaaaar	tgaccttgr	gcaattyat	aaagaataaaa	tatttctagt	tttttgggtgc	600
	tgaactgcta	aaagatggtt	ctatacatgt	aacagggtggc	tttagttgg	ttgctttcac	660
	tgaaatttga	ttcaaataaa	gcattgcatt	attttacctt	tggaaattata	aaaaaaaaaa	720
45	aaaaaaaaaa	aaaaggccgg	cc				742
	<210> 18						
	<211> 1219						
	<212> DNA						
50	<213> Homo sapiens						
	<400> 18						
	aacgcactca	atattcagaa	gtttgaattt	taccactctc	aaacacagtt	aaaaagatag	60
	ctgtttgaga	atgcttctt	actaatacta	gtacaatatac	ttaataatgt	tatgtacatt	120
55	atagaaaatc	ttgaacagta	caagattttc	ataattaagg	catgcaaaac	tgcttgggt	180
	ctttgattcc	aggtgtcctc	ttctcccttc	tgcttttgc	atctatgtt	aatataattc	240
	taacccagtc	taagtatgg	aaaaatttctt	accctgcctg	cttttatagc	tcatcaaatt	300
	tccctgtatc	agctatcaact	tttctggtag	gtgttagtctg	atttctgtct	gtcatgcctt	360
	tgccacaatc	cttttttga	agagtaggtt	aaagatctat	taaagtgtt	atcacattgc	420
60	ttaaatatat	aaagcctcca	gtggtttccc	atatcaactct	gtaaaaatgc	ccttgcac	480

	ctctcccatc aacctcgctt tttctgtct tggatatgca catctcttcc tgagcctta	540
	ttgcacatcct catgtggga tgggtctgtc tcagagatag tctttattca ggtcccactc	600
	tgcagtccct tcacaggggg ctgccttcac cacccttct aagtaaggct ctctaaacac	660
5	ctctatcata ttctatccct tagccagcac taatttttc ataatgctta ccactaactg	720
	aaatttactt tattcattaa tctcttcctc attagaatgt aagctcgtaa gggagggca	780
	gctctgttagt ttattcatta ttgttatgtcc ctcacctaatt cctatgagtg tctggccat	840
	attagggtat gtaataaaaata ttacttgagg aatgaatgaa ttaacatac tacaattct	900
	ctgagtgact cttttttaag ctttcatacat cattcacact ttcttgctt tcataatggc	960
10	atgtccaatc acccttccat gaatatctgt acctgttaca aagagaggac taggttcctg	1020
	gagttcatag atgtaaacaa catccttaggg ktagcaaact ggtggccct gagccaaatc	1080
	ctggcctgca catgtatccc gtttgakttt tacaatgtt gtataaaatg aactggctga	1140
	taatttttt taattggaaa tgtttacatt aaaaacctar acttctagct gctcttaaaa	1200
	aataaaaaata cggtgggc	1219
15	<210> 19	
	<211> 874	
	<212> DNA	
	<213> Homo sapiens	
20	<220>	
	<221> SITE	
	<222> (461)	
	<223> n equals a,t,g, or c	
25	<400> 19	
	ggtcgaccca cgcgtccgag caattgaatc atctgccaa ggataagctg ctggtgagag	60
	cagagttggg atttgaagtc gagttagacc ccagtgtca cagtttgac gattaaattc	120
	ttccagctt catttttcac ttagataatg gtatgtatag tactgaccc taatgtgtgc	180
30	atttgtgggt atgtggtcca ttcaagctta atccccagaa gacaaggctt attcctttc	240
	ttatttttgg tcatgtttta tttttccatt gcttttaaca ggattaccaa aggcacactc	300
	agtagtcagt aaacacattt cttagaaagg tgggtgtca tcatgccaca tattcatact	360
	ttcctgggtt gaaaaataga tcatcagtaa aaacatacag gaaaaatgaa tcttgccaat	420
	gcaattgtta acctacaacc ataataatacc ttaagtatat nttgcacat aagtataaca	480
35	tgcgattnaa aacaataaaac cagattgaga tctaaggagc atttttaag taattactaa	540
	tgtttatccc agagagatca cacaacttca aataaaaact gacatgatgg aacacccctg	600
	agaataaaact ttagtgc当地 atggaaaata atttttaca agtaaatttgg aagaacaatg	660
	tgaactttct ataattatatac agatcata tactgatgg caaaaatgag taattttgc	720
	atattaatata ttcacttata agaatgcata ccacctgatc caggatggg tccaggaaca	780
40	gaaaaagaac attagktaaa aatgacagaa atctgaatat agtataagat agctaaaaac	840
	aaaccaaaaa aaaaaaaaaa aaaaaaggc ggcc	874
45	<210> 20	
	<211> 464	
	<212> DNA	
	<213> Homo sapiens	
50	<220>	
	<221> SITE	
	<222> (21)	
	<223> n equals a,t,g, or c	
55	<400> 20	
	caaacccttc agtggatgag nccaagtcgc agaaagcatt ctgttgacag atgaacagcc	60
	gaaagctggc cagaccctcc tgkatgact cccwgccck tktatcagaa acacaggcaa	120
	ggaaatttggg actgcccacc agcccacat ggtggctaa ttgggtgggt gcgttgc当地	180
	ttgtctcttc gttttgttaa gtttttaat aagtagtgg tgcataatgt cttttatgg	240
	gtttgtataa ttgttaacgg ttttagcagc ctataactt tcaatgtgg cttttactta	300
60	ggaaaaaaa caatttgtaa atacagaaca ttgtttaaaa gacataacca tagaacatag	360

	cttcctgttt gtggattttg tttcctatat attcaaagta aaatgactta caggaaaaaa	420
	aaaaaaaaaaa aaaaaaaaaaa aaaaaaatcg gggggggggc ccgg	464
5	<210> 21 <211> 637 <212> DNA <213> Homo sapiens	
10	<400> 21 gattttcctg cttgcatcat ttcttagcaca gagctggagg aaatggcgag gtgcagggtgg ccgctggccm tgctttcta catgggagca agacagctgc tagtgaagg ggaatgacca ggcagccaca gggaggacat gtggcctcag gaaggctggg tgttatcct gtttctgcta ggaacacgtg tggggcttg tgggggtgac tctctggctc cccaagcctc cttttctac	60 120 180 240
15	tgttatatcc taaaagtgcc tctgaggcca aagcctttgt ggcaattgtc aaatgagtcc atatgcagtg agtacctgt tgagggagga caaggtcacc aagagctgag aatgtttctc cgactgtatgac gacctagata ttgggtacat ggaggtcccc ggtccctttg tgattcctgc agcctgttc ctcctgcct ggacccccc tcagctcaga aagccaaattc cctagattcc	300 360 420 480
20	aaaggccttc ccagaccaat tagcatgtcc tgcagctgtc agctccctgt gcctagcctg gacctcagct catgtctagc acccagtctc ccaaccccac acatattcac aaataaaaaga aaataacaaa tgaaaaaaaaaaa aaaaaaaaaaaa aaaaaat	540 600 637
25	<210> 22 <211> 752 <212> DNA <213> Homo sapiens	
30	<400> 22 gaattcggca cgaggggatt acaggcatga gccaccatgc cggccatat aaagcattta ggatagttag ttgctatTTT tattttttt ttattgttgt tgttattata ttactacttt atcccatttc acaaggatgg catgttgcca acattgtctt tctaaagaat atctctgatc acatccttgt tctattaaaa accttttggaa agctccctct tacctttaga agaaaattggaa	60 120 180 240
35	acttcatgtat tcctcatgtt ctggctccag cactgagttc ggaatgctag tgtgagatga ggccttagaa gtcatccagc tgaactcctg gaatttttat agatgaaataa attagcatc cagacattt tcytggc cccctgtamg ccatgtccctc ttccagactc ctggataaga ctgrcagaca tcaccattct cttaaaccag aactacactt gccttcatcc attgatcac ctgggtccag gtaactcatg agctttgttag ctcccttct ctcagacctt ccaaggaaga	300 360 420 480 540
40	caatggcata attttccca tatgtctaa ttagcaacct ttccctgccc ttctgtgggt gggcaggggc ggacacagtg ggtcacacct gcaacctgtt atcccagcac ttggggaggg tgaggtgggc agattgcctg agtcaggag ttcaagacag tctggtaac atggcaaaat cctgtctcaa aaaaaaaaaaaa aaaaaaaaaactc ga	600 660 720 752
45	<210> 23 <211> 492 <212> DNA <213> Homo sapiens	
50	<220> <221> SITE <222> (486) <223> n equals a,t,g, or c	
55	<400> 23 aagctggact cgcgcccttg caggtcgaca ctatggatc cmaaagaatt cggcacgagc aaggacccag aagttagggtt ttggcctagg taacggggca gagatgtgt tcgagattct ccccggactc tccgtcatgg gcgtgtgtt gttgattcca ggactggcta ctgcgtacat ccacaggttc actaacgggg gcaaggaaaa aagggttgc tattttgggt atactggag	60 120 180 240
60	tctgatggaa agagatggc gcatctctgg agttgatcgt tactatgtgt caaagggttt	300



5           <210> 26  
 <211> 917  
 <212> DNA  
 <213> Homo sapiens

10          <220>  
 <221> SITE  
 <222> (434)  
 <223> n equals a,t,g, or c

15          <400> 26  
 gaattcggca cgagggttca ttgcctcaa cattctctgt ttttcaccga atcatgtctg         60  
 tttcctccaa cctctggcaa acactgatct tgtaactgtc ttgtgggtt tgccttttc  
 cagaatgtca tatagttgaa atcatacaat tgcgcagact ttttagattt cttctttca  
 cttagtaaca tttaagttt ctccacccct tttcatggct tgatagttca tttctttaa  
 ttgctcaata ataaatattt cattatctag atagaacggt ttatctacct agtgaaggac  
 atctcaattt cttccaagtt taggcaaata taaacaaagc tgctatcagg attttcaca         360  
 20          gaggaaaaga cagtggatc caaaaactgaa tggtctatca ataaatgacg catggatcat  
 ctacacccat grancattt tgcatccatg agaaaaaatcc agatgttaga aggttatgtat  
 aattttcgag aaaagagt gtaactggaa acaccaargaa aaaaaggaaa tggatctata  
 tatttaggtt gagatattt tggctgcga gaagaaaatat attattttt atactagata         420  
 gttaatgtt gccttggtg gcaagaaaag gtaaaaaggc agaaggaggc ccaacccaaa         480  
 25          gaggaagagg aaaaaaaaaa aactgcacta agaaaaatct tttaaaatgtc tggatcaca  
 gccagggtgca gtggctgaca aatgtaatcc cagctacttg ggaggctgag gcaggagaat         540  
 cgcttgaacc caggaggctg aagttgcagt gagctgagat catgccattt cactccagcc         840  
 tggtgcacaga gactctgttt caaaaaaaaaa aaaaaggat atgatcacat ctgtgttaac         900  
 ttacagacta gtctcgaa   917

30

35          <210> 27  
 <211> 662  
 <212> DNA  
 <213> Homo sapiens

40          <400> 27  
 gaattcggca cgagggttccc atggcacttt atatgtgtgc atagagagcc agggaggcagt         60  
 ggggttcagg gtggggccat gctatgtgct gcagagctgg tgggtcacag tctccccagg  
 tgatggtgtt gttaataatc atcctaggcc cgtgggggtgg ggtgaggatt gatgcattgag         120  
 aaagttgagg cggggggccct ggcatggagc agggctcagg ccgctgtca cccaggctca  
 tgcagccct ccggagccctg tgggtgtata ggggaagcgc aggggttctt cagccagagg         180  
 gacaggttca rggcctgctg atgcccctt ctggttttgg gaccttgagc aagtccctt  
 gccttttgtt gctgtgcctc ggtttcttct tctataagaa ggaggtgatg atgtaaccca         240  
 45          420  
 cccacccaggc ccctctaccc cgcgcattcag ggtgcaggc gagctagcac tgcgtgcacca  
 ggagtgaggc tggccctgg cggggccacgc ctggagaggc atgcctatct ctgctgcccc         480  
 cctgtggcgt catcatatca acctgcccagt ccccccattt tgggtgttaat ctcccaaggagg  
 atggggactg rttctgcata ttctttgcta aacaaagacg ctgtttggc tggctctc  
 ga   540  
 50

55          <210> 28  
 <211> 699  
 <212> DNA  
 <213> Homo sapiens

60          <400> 28  
 gattcggcac gagaaacttt taaatcttta gttatcttctt aataactttaga acactttttt  
 aaaaactttac aaaacaaaag agcagaataa ttagatctt tcaggagaat atgacttttt  
 tttcttaaagc acactggacc atagaggaag accaaaggaa tgtacagttt cctgtctccctt         60  
 120  
 180

	cctgacttgc	tgtatTTGAC	tctgtccccca	ctgggtggTGG	caatgtatt	aacCCCaCac	240
	tttaacgtgg	caaATCCCca	gaatCTGTTG	gCTGGTCTCT	ggctAGAGAA	TGAGCACAGT	300
	ttcacCCta	tggCTCCAGA	aAGAGCAAGA	ACACACCCT	gccAGCCAGA	AGAGAGAAA	360
5	gtcttGTTCT	gtCTCTTCC	CATTGTCCTA	aatAGCCAAg	cACAGGTTCa	ACCACCCCA	420
	atGCCACCCt	tCTGCTGTGc	AGCAGCCAAG	GAAAAGACCC	aggAGGAGCA	gCTCCAAGA	480
	cctCTGGC	gtCAGTGC	CC	CCAAATTCTT	TGTGTC	CCACACTCG	540
	ctgacaaaag	ccaACACTT	gtCTCTTCTT	TTTTTTTTT	CTTTTTTTT	GAGCAGAGTT	600
	tcactCTTGT	cACCCAGGCT	GGAGTGCAAT	GGCAGGATCT	TGGCTCAT	CAACCTCCAC	660
10	ctcccgggtt	caAGCAATTc	TCCTGTCTCA	GCCTCTCGA			699
	<210>	29					
	<211>	1637					
	<212>	DNA					
15	<213>	Homo sapiens					
	<220>						
	<221>	SITE					
	<222>	(726)					
20	<223>	n equals a,t,g, or c					
	<220>						
	<221>	SITE					
	<222>	(727)					
25	<223>	n equals a,t,g, or c					
	<220>						
	<221>	SITE					
	<222>	(728)					
30	<223>	n equals a,t,g, or c					
	<220>						
	<221>	SITE					
	<222>	(899)					
35	<223>	n equals a,t,g, or c					
	<220>						
	<221>	SITE					
	<222>	(901)					
40	<223>	n equals a,t,g, or c					
	<400>	29					
	aaatgtGCCA	cgtCTTCTAA	gaAGGGGGAG	TCTGAACTT	gtCTGAAGCC	CTTGTCCGTA	60
	agCCTTGAAC	TACGTTCTTA	aatCTATGAA	gtCGAGGGAC	CTTTCGCTGC	TTTTTAGGG	120
	actTCTTCC	ttGCTTCAGC	aacATGAGGC	TTTCTTGTG	gaACGCGTGC	TTGACTCTGT	180
	tcgtCactTC	TTTGATTGGG	gCTTTGATCC	CTGAACCAGA	AGTAAAATT	GAAGTTCTCC	240
	agaAGCCATT	CATCTGCCAT	CGCAAGACCA	AAGGARGGG	TTTGATGTTG	GTCCACTATG	300
	aaggCTACTT	AGAAAAGGAC	GGCTCCCTAT	TTCACTCCAC	TCACAAACAT	AACAATGGTC	360
	agCCCATTG	TTTACCCCTG	GGCATCCTGG	AGGCTCTAA	AGGTTGGAC	CAGGGCTTGA	420
	aaggAATGTG	TGTAGGAGAG	AAAGAGAAAGC	TCACTCATTCC	TCCTGCTCTG	GGCTATGGAA	480
	agaAGGAAA	AGGTAAAATT	CCCCCAGAAA	GTACACTGAT	ATTTAATATT	GATCTCCTGG	540
	agATTcGAAA	TGGACCAAGA	TCCCATGAAT	CATTCCAAGA	AAATGGATCTT	AAATGATGACT	600
	gGAAACTCTC	TAAGATGAG	GTTAAAGCAT	ATTTAAAGAA	GGAGTTGAA	AAACATGGTG	660
	CGGTGGTGA	TGAAAGTCAT	CATGATGCTT	TGGTGGAGGA	TATTTTGAT	AAAGAAGATG	720
50	aAGACNNNTA	TGGGTTATA	TCTGCCAGAG	AATTACATA	TAACACAGAT	GAGTTATAGA	780
	gatacatcta	CCCTTTAAT	ATAGCACTCA	TCTTCAAGA	GAGGGCAGTC	ATCTTAAAG	840
	ACATTTAT	TTTATACAA	TGTTCTTCT	TGCTTTGTTT	AATTATTTT	ATATATTNT	900
	NCTGACTCCT	ATTTAAAGAA	CCCTCTAGGT	TTCTAAGTAC	CCATTCTTT	CTGATAAGTT	960
	ATTGGGAAGA	AAAAGCTAA	TGGTCTTGA	ATAGAAGACT	TCTGGACAA	TTTCACTTT	1020
55	CACAGATATG	AAAGCTTGT	TTACTTTCTC	ACTTATAAAAT	TAAAATGTT	GCAACTGGGA	1080

	atataaccacg acatgagacc aggttatagc acaaatttagc acccttatatt tctgcttccc	1140
	tctattttct ccaagttaga ggtcaacatt tgaaaagcct tttgcaatag cccaggctt	1200
	gctatttca tggtaataatg aaatagttt tggtaactg gctctgagtc tctgcttgag	1260
5	gaccagagga aaatgggttggac ttgttaatgg ctactgctt actaaggaga	1320
	tgtgcaatgc tgaagttaga aacaaggta atagccaggc atggtggtc atgcctgtaa	1380
	tcccagcact ttgggaggct gagggcgccg gatcacctga gttgggagt tcgagaccag	1440
	cctgaccaac acggagaaac cctatctcta ctaaaaatac aaaagttagcc gggcggttg	1500
	atgcgtgcct gtaatcccag ctacccagga aggctgaggc ggcagaatca cttgaaccccg	1560
10	gaggcgaggg ttgcgttaag ccgagatcac ctccagcctg gacactctgt ctcgaaaaaaaaa	1620
	aaaaaaaaaaaa aactcgaa	1637

	<210> 30	
	<211> 2142	
15	<212> DNA	
	<213> Homo sapiens	
	<400> 30	
	aattcggcac agagacgagg gtccccgggt ctgacaggag cagcctgtgg gcacccggc	60
20	ggtagttgga ggcggagag ggtccgtage cgccggccccc tgcccccca tggccctcct	120
	gtcggaccgg gttcggccgc gcgcgcgtcgc ccgcctagtg ctgcgcctca acgcggcggt	180
	gtgcgtgcgtc agctacgtgg cggcgcgtc ctggttcttg ggcgtgttt tccccccgct	240
	gaccgcgcgtc acttacatgt cggagaacgc catggcgtcc accatggtgg aggacgtt	300
	tgccggccga gaccgtggcc gggctttgc ccgggacttc gcccggccacc gcaagaagtc	360
25	gggggctctg ccagtggcct ggcttgaacg gacgatgcgg tcaatgggc tggaggtcta	420
	cacgcagagt ttctcccgga aactgcctt ccacatggag acccacagc gctatatggt	480
	gtcgggcacc aacgtgtacg gcacccgtcg gggccggsgt gtcgcacca cccggcgtcgt	540
	tgtgtcacc gtgcctgtg gctctgactc taccaacagc caggctgtgg ggctgctgt	600
30	ggcactggct gcccacttcc gggggcagat ttatggggc aaagatatcg tcttcctgg	660
	aacagaacat gacccatcg gcaactggc ttggcttggaa gcctaccacg atgtcaatgt	720
	cactggcatg cagtcgtcctc ccctgcaggc ccgagctggg gccattcagg cagccgtggc	780
	cctggcatcg acaatgtatc tggcaccagg cctcgatgtg gccgtggagg ggcttaacgg	840
	gcagctgccc aacccatcg tgcgtcaatct cttccatggc ttctgcctca aagggggcct	900
35	gttgcacccg cttcaggccgca agctgcaccc cgaggactgg acatcattgg atggaccgt	960
	gcaggccctg cagacactgc tgcgtcatgg tctgcggcag gcctccggcc gccccccacgg	1020
	ctcccatggc ctcttcctgc gtcaccgtgt ggaggcccta accctgcgtg gcatcaatag	1080
	cttccgcctg tacaagtatc acctgggtggc agtggggcaag gctttggagg gcatgttcc	1140
	caagctcaac cacccctgg agccgcgtca ccagtccttc ttctctact tgctccccgg	1200
	cctctccccc ttcgtctcca tggcctcta catgcccgtc gtggcttct tgctccctgg	1260
40	ccttggtctc aaggctctgg aactgtggat gcagctgcat gaggctggaa tggcccttga	1320
	ggagccccc ggtgcctctg gccccagggt accccctccc ccacacagg gtgtggggct	1380
	ggccctcgctc gtggcacctc tgcgtatctc acaggccatg ggactggccc tctatgtcct	1440
	gccaatgtcg ggccaaacacg ttgcacccca gcaactccca gtggcagagg ctgaggctgt	1500
	ggtgctgaca ctgcgtggca ttatgcagc tggcctggcc ctgccccaca atacccaccc	1560
45	ggtggttaacg acacaggccc cagacagggg ctggatggca ctgaagctgg tagccctgtat	1620
	ctacccatcg ctgcacgtgg gtcgtatcg cttccatggc ttctactgg gcttcctgt	1680
	ggccacccacc atgggtccca ctgctgcgtc tgccaaaggct catggggccc ggaccctcta	1740
	tgctgcctcg ctgggtctga ccagccggc agccacgc cttggcagcc tggccctgt	1800
	gcggggatcg caggaggccg cactgtcact ggccggaggc tggcagctct tccctggcage	1860
50	gctagccac ggtgtgtctgg agcaccacac ctacccgcctc ctgtttccactgtcgtc	1920
	cctggccctc taccctgtct ggctgtttt ctggatgtg ctcttcgttgg aatgtgatct	1980
	gcctgtccgg gctggggacag agactccccca aggacccat tctgcctct tctggggaaa	2040
	taaatgatgt tctgtttcag carmaaaaaaa aaaaaaaaaa aaaaaaaaaa aatgaccctc	2100
	gagggggggc ccgggtaccc aattggccct atgaagagggc ga	2142

55	<210> 31	
	<211> 1564	
	<212> DNA	
60	<213> Homo sapiens	



	aaatgcttt aaatgctctc tcacatcttgc ttcccccact accccccact ctttaggtatg tatgatgcata atcttgccttca taagtaagtt tcttcctgtct ccttttgat ctccctttct tgtctttctt cctacccttt gtctcttgggt gttttggac tttttttttt ttttttgcc ttttgtacaa agattagttt caatgttagtc tgttagcctcc ttgtaaaacc aattaaaaag 5 ttttttaata a	1440 1500 1560 1620 1631
	<210> 33	
10	<211> 978	
	<212> DNA	
	<213> Homo sapiens	
	<220>	
15	<221> SITE	
	<222> (2)	
	<223> n equals a,t,g, or c	
	<220>	
20	<221> SITE	
	<222> (27)	
	<223> n equals a,t,g, or c	
	<400> 33	
25	angagttgca tgcaagsgta agttgggnccc ytsgrggatc ttttagagcg ggccccc ttttttttt tgcatgtctg agtttgcga ataagattca tatttactac aagtaatgg attggaggat cagaggggag aaagtcaactt atcacattag tgtaatttc tgatggtagg attatggaga gtttttagttt tttccctttt ttccccaact tctctccccct cagatttta aaaataaacaat tgcgtgggtt gttttttttt gtttttgatt gttttttgtt ttcaaacag gtctcactcc tatccatgtt ggctagatgt cagtagtgca atcttggctc actgcagcct 30 cgacttcctt agctcaggtt atcttccac ctcagtcctt tgagtagctg ggactccagg tgtgtgccac catgcctggc taaatttttt tattttttt agatacaggg ttcaccatg ttggccagac tggcttgaa ctccctggcc tgcccaccc acgtctccaa agtgcata ttacaggcat gtgccaccat atccagccta ataacattgt ttttaatgtt cattaagtca tcccacccctc tcagtttgc agaagccctc caagagggac agaatacgat gcaaagtacc 35 atttctgacc ctgagacatg gatattttt gttcatatgg atgtcacctg aaaaacccac tcactcaaat ggtctgtgaa gcttgcaaaa acaggaatgc ttacccctct gggtcctgaa tttttgggttc tcttgactc ttggaaattt ttctttctca gaaaggagcc ctctttctat ttccccccttca agttgtgact tgaccctcac atcccttttct tctccaggcc cccttgata 40 gatttttttta aaattttttt ggaggccatc ctttttagga agacggacgc gtgggtcgac cgggaaattcc ggacgta	60 120 180 240 300 360 420 480 540 600 660 720 780 840 900 960 978
	<210> 34	
45	<211> 898	
	<212> DNA	
	<213> Homo sapiens	
	<220>	
50	<221> SITE	
	<222> (402)	
	<223> n equals a,t,g, or c	
	<220>	
55	<221> SITE	
	<222> (452)	
	<223> n equals a,t,g, or c	
	<400> 34	
60	gaattcggca cgagattatg tagtagatgt cactagaatt cttgaaattt gtcttcaagt catggcagta tttcagttatc gtcctttgg gattgcctga gtgatactca agagtttagac	60 120

	tagttttatc tgggttcttt	gaagaaccgg	ggacacctca	ctggcttatg	ttgaatttct	180
	gcactgcagg	gaccaactat	aaatgggttt	tttggtttt	tacgtgttaa	240
	atgttaattct	tccatcatt	catgcacaaa	tgttotcaca	caaattgctt	300
5	taaaaacttg	aataattttt	ccctgaagaa	atgttgaact	tttctgcaag	360
	kggagcgcgt	gttggaaaggc	ytgaakggg	ccgtactgta	cngcctawtt	420
	aaattawgat	ttcyattttt	watycattta	cngatgactg	aatakgyca	480
	tatccccta	tttcaaaatg	cagcaatcta	taaacaaaat	acttgcatt	540
	acacccctt	ctataatttg	tatagaaaaat	taagtgcag	ggccaggcac	600
10	ctgttaatccc	agcactttgg	gaggccaagg	cgggtggatc	gcctgaggtc	660
	gaccacccctg	gccaacatgg	cgaaactcca	tctctactaa	aaataaaaaa	720
	ggtgtgggttgg	cagacgcctg	taatcccagc	tactttggag	gctgaggcat	780
	tgaaccagg	aggcagaggt	ggcagtgagc	ttagatggcg	ccattgcact	840
	taacaagagt	gaaaactgaa	gctgtctcaa	aaaaaaaaaa	aaaaaaaaaa	898
15						
	<210> 35					
	<211> 754					
	<212> DNA					
	<213> Homo sapiens					
20						
	<220>					
	<221> SITE					
	<222> (311)					
	<223> n equals a,t,g, or c					
25						
	<400> 35					
	cagccatc	tcctgttggc	cccttgtatg	taccctgtgt	ttgagttgta	60
	gcttgccat	aatctttctt	ttaactcctg	tgcttccttc	tcatcctttg	120
	actttctgct	taaagtggac	cttgacttct	ctttatcttg	ctccatttgc	180
30	tgtcccaac	tgcagtgcta	attccttgg	aatgtttat	aactttgtca	240
	cactgtaagt	atagaacatg	ctgggaaatc	caaattaaaa	atgacagttg	300
	gacttctggg	nagggaccaa	ggaaaatgt	ccagagtggc	aggatagctg	360
	ggattgccag	caatgtaaag	cgtagactcc	agaggaacag	tgtaactta	420
	gcaggcatca	gtacttctgg	ttctgatggc	ccggggattt	ctaagtagta	480
35	gcattatttg	ttatacagtc	tactgctaga	tgaacaaggc	taagtctaca	540
	attatagaaa	ttaggcccccg	tctctgttaa	gaataaaaaa	aattagccgg	600
	ggggctctgt	ggtcccagct	actcgggagg	tgacccagga	gaatggcg	660
	gccccggctg	cgtggggccg	agatagcgcc	actcgagtct	ggcctggcg	720
	actccgttctt	aaaaaaaaaa	aaaaaaaaact	cgta		754
40						
	<210> 36					
	<211> 699					
	<212> DNA					
45	<213> Homo sapiens					
	<220>					
	<221> SITE					
	<222> (483)					
50	<223> n equals a,t,g, or c					
	<400> 36					
	gaattcggca	cgagcggcac	gagccacctt	ctcagtccag	tctatggta	60
	tctgctaaa	acccatcctt	gcttctttgt	tgccttaccag	atgcaggctcg	120
55	cctccttccc	ggactcagga	acagcaagac	tgttactatg	ccattgtccc	180
	tcccacccctc	cttttttttc	cctctccccc	tcccttcttt	caccccttcc	240
	atgcgttcc	aagtattaat	tttaaaaattg	ttctacaaga	atgcgattta	300
	gtgaaccaag	cagaatttct	tagtatttt	ttgccttagg	gcattccct	360
	aaaatttgc	ccccattcct	ttttgcctgt	ggaacttate	tttattcttc	420
60	tamtcttaat	agcactttga	attnaacctc	cctggtagtt	tttctcagcc	480

	ttnctgaaaa caggattctc tggctccat gtctggctaa ttttgtatt ttttgtggag	540
	acaaagtctc actatgtgc ccaggcaggc ctcaaaccacc tggcctaag ccatectccc	600
	accttggct cccaagtgc gggattataa gcatgtgcc a tggacccag ccagagaccc	660
	tgtctttta aaaaaaaaaaaa aaaaaaaaaaaa aaactcgta	699
5		
	<210> 37	
	<211> 971	
	<212> DNA	
10	<213> Homo sapiens	
	<400> 37	
	gccaccggc cgca gttccctt gggtcgccg gcagctgtga gcccggagg caaggcggtg	60
	cagaccgcca tcctggccgg cggcatgagc gtgggtcgcc cctgcgtgtc cctgacccag	120
15	tgcctcaggc atctggcgca accccgacgg ggcgcacaaga tgctcgacca cagggagagg	180
	ctgaggaact cggcctgcgc cgtgtctgaa ggctgcaccc tgctatctca ggctttaagg	240
	gagaggtttt cggccaggac ttaccgcca gtgaattcca attctgtgaa ttagcaccccc	300
	accccccatac cccttcttcc acccccagac taaaaggaga tacttactct ctgccccctct	360
	ccatttatac caaaggaaatc ataggtaaaa ccccttaccc tcccaacgt taaatgtcg	420
20	agaggaatct tccacaaggc agggccatgc acgcaacctg cacacgcact tggagggccc	480
	aggtgtctctt ccaccagccc ccatcgacca gggactggaa gatatgtcat ctgctggttt	540
	tgttatcaact cccacccctt acccccagccc gtsttccgga atttctcaac taaaatttsat	600
	tattggcag gaaggaggc atgggttcat ttcatttttgg tttttgtgt ttttaattaa	660
	aagaaagggtt acctcagttt tcaactccta gacatggatg tagctaccc tttttgtatg	720
25	tctttttttt tttaagcaat cgtgttgaat taggatata cttgggtgtgg aaagagtatg	780
	aatttgcacat gtgatttgc aatggggggg agctactgtg agegtgtgtt ttttaattt	840
	acactataga gtgattttttt ttccccca cgtcaagttt ttaccttgca tttactggag	900
	tatttatttc atctattaaa atgttatgtt ttcagaaaa aaaaaaaaaaaa aaaaaaaaaaaa	960
	aaaaaaaaactcg a	971
30		
	<210> 38	
	<211> 872	
	<212> DNA	
35	<213> Homo sapiens	
	<220>	
	<221> SITE	
	<222> (2)	
40	<223> n equals a,t,g, or c	
	<400> 38	
	tngcgttct ccacaccgaa gaggacggtg ggcgcacaaca gacaggcgat taatgcggct	60
	cttaccggc caaccaggac tacagtatac attgtggaca ttcaggacat agattctgca	120
45	gctggggccc gacctcactc ctacctcgat gcctactttt tcttcccaa tgggtcagcc	180
	ctgaccyttt atgagcttag tggatgtatc cgaaatgatc aggactcgct gatgcagctg	240
	ctgcagctgg ggctgtgtgtt gctgggtctcc caggagagcc aggagtcaga cctgtcgaaa	300
	cagctcatca gtgtcatcat aggattggga gtggcttgc tgctggctt tggatcatg	360
	accatggct tcgtgtgtgt gccggaaagagc tacaaccgga agcttcaagc tatgaaggct	420
50	gccaaggagg ccaggaagac agcagcagggt gtgatgcctt cagccctgc catcccagg	480
	actaacatgt acaacactga gcgagccaa cccatcgta acctccccaa caaagacactg	540
	ggcttggagt acctctctcc ctccaaatgac ytggactctg tcagcgtaa cttccctggac	600
	gacaactctg tggatgtggaa caagaacagt caggaaatca aggacacag gccaccacac	660
	acaccaccag agccagatcc agagccctg agcgtggtcc tggatggacg gcaggcaggc	720
55	gcaagtggac agctggaggg gccatccat accaacgctg gctggacac cacggacact	780
	tgacaggggc cccacttctt ctggacccct tgaagaggcc ctaccacacc ctaactgcac	840
	ctgtctccctt ggagatgaaa atatatgacg ct	872
60	<210> 39	

<211> 608  
<212> DNA  
<213> Homo sapiens

5 <220>  
<221> SITE  
<222> (10)  
<223> n equals a,t,g, or c

10 <220>  
<221> SITE  
<222> (16)  
<223> n equals a,t,g, or c

15 <400> 39  
ccataacgcan accgcntctc cccgcgcgtt ggccgattct tatggcagct ggcacgacag 60  
gtttcccgat gaaaaaggccc cagtggcgc aacgcatta atgtgagttt gctcaactcat 120  
taggcaccccc ggctttacac tttatgcctc cggctcgat gtktgttgg atttgtgagcg 180  
gataacaatt tcacacagga aacagctatg accatgattt acgccaagct cgaaattaac 240  
20 cctcaactaaa gggAACAAAAA gctggagtc cacggcggtt cggccgctct agaactatgt 300  
gatccccccgg gctgcaggaa ttccggcacga gtttgggtgg agtttccaag gtgaaagttt 360  
ctgaatttgtt caatcagtga cgcctttgtt aagatggctc atgtgggtgtt cgctcgcaat 420  
gaatgcctga taagggcttt tctgtttttt ttgcactgtt taagtttgc cccatcgcc 480  
25 ggggaaagtta atatcagaca cacactttt acggtagaaag agagggttgac tactccaagg 540  
gcactgaaac tctcaactgag ccttattgtt tctctacacg cgamtgcag aaagcaggag 600  
tgctcgta 608

30 <210> 40  
<211> 855  
<212> DNA  
<213> Homo sapiens

35 <220>  
<221> SITE  
<222> (850)  
<223> n equals a,t,g, or c

40 <220>  
<221> SITE  
<222> (851)  
<223> n equals a,t,g, or c

45 <400> 40  
ctgttaatagc acacaactca gaactcttca gcatttgtt gattccttac ctctggctga 60  
taaaactcta atgggttgtt gtttactttt tttccatttt ctttggcttt gtgcaatttt 120  
tgtgttaactt tacttgtacc tatattttt gtttacagtt ctttttaagg ggaggggttag 180  
gttcttaaga tcttgttgtt tattgttagat aaaaattttt tctgttgttta gaaaagcatg 240  
gtttagtcgt ttgactgaaa aagacactgt attattttacc aaaggggtat tgtttttgca 300  
50 tttgtttata aatgcattat ttttgtactg taaaatttggaa cataatttct gagtttatta 360  
ctactggcat tttctttttc cttttttttt ttttttaacc gtaagtgcac gatgcagggt 420  
cataggcccc agaccaaaact agaccacccag catgttcatg tccagacccgc ggcagtggcg 480  
tgcactgctt gtgcacccca gttcctccag ttttgtttt aattcagcat 540  
55 cctgtgtgtt ttactttcca agcaagatct gttgcactc ccaaatgcgt tttaatgagc 600  
tcatctttat ttgcctttct tcttacgtat ttttgttattt agattgttca ggagatattc 660  
tagaaggcat taatggttt cattcaaaaac gatgtggttt gtccaaatgttta ttttctgtct 720  
ttattactga gacggattaa tctccttattt tttttcttga tgatttgaag ttgttaacagt 780  
tgtccagcta ttgcttaata aaattttgc gatcaaaaaaa aaaaaaaaaa amctcgcccc 840  
ggccccccgn nccca 855

5       <210> 41  
 <211> 1042  
 <212> DNA  
 <213> Homo sapiens

10	acggcccgta attccccgggt cgayccacgc gkccgtgctt cctagaaggt cgtgtcacgt	60
	ggaacctctt aatctcagca tccggagctc caggaaggga aaatttcaag tcagatagaa	120
	ttctatatat accatttctt tggAACCTTC agccctcaag attccaacat catgaccta	180
	gtttcaacac agttgcctt agtcctcatg tcactgcttt tggtgctcc tggttgaa	240
	gcagtagaaag ccggatgc aatcgccctt ttgttaggt tggttctcg cattacaggc	300
	atttgtgcct gcttgggggt atatgcacga aaaagaaaatg gacagatgtg actttgaaag	360
	gcctactgag tcaaaccctca ccctaaaaac ctttgcgtt tagaggctaa acctgagmtt	420
15	tgggtgtgtga aaggccccaa gaatcgataa ataagggagt ttcacattt tcattgtttc	480
	catgaaatgg caacaacat acattataa attgaaaaaa aaatgtttt tttacaacaa	540
	ataatgcaca gaaaaatgca gcctataatt tgcttagttt gtatcaaag aagtagatg	600
	gctgaaattt acataagtaa tatttcataa tctttagaatt ctctcaaaac atgtgaaata	660
20	ggaagaagga agttctgcc cagaatctta gaaatcacc actgttccgt tataatcact	720
	gcctcctgaa tcgttgagga gtctttaaa ttagatttt gttttgtgt ctcccaagtt	780
	aatatttatat tttagatcatc gagatcagg yaaaaaggaa aacttttac tctagggaaa	840
	aaacatttag aaaaatgtat tcagtgtatc taataactgaa atgcggaaaa aaatttaatg	900
	ttaaaaaaaaaa actatagaca ttgacatgga aaagagattt aatgttttga aaaaaaaaaactt	960
25	tatattaact gagtaacatc ctccgtatga gaagtactat attaaatata aaccattat	1020
	gttataagtt aaaaaaaaaa tt	1042
30	<210> 42 <211> 702 <212> DNA <213> Homo sapiens	
35	<220> <221> SITE <222> (515) <223> n equals a,t,g, or c	
40	<220> <221> SITE <222> (614) <223> n equals a,t,g, or c	
45	<220> <221> SITE <222> (673) <223> n equals a,t,g, or c	
50	<220> <221> SITE <222> (677) <223> n equals a,t,g, or c	
55	<400> 42 gggacaatga actccttctg gtctaagtta ttgggtgctgc ccctgctggc tccgctgtcc	60
	atggcccgag cctctgcctg tcagagatgg tagagccacc aggacatgga gtcattgctg	120
	acacaggaa acatgagatg tcttaggttt ggtgtatgtg aaacatgcat gagaataga	180
	ggccaaaagt tccactgtgg agcgcagaca gaatggctcg aatgcttgc cagttactac	240
	gtcagtagtt tgcataatata tataattat acatctataa cctatgtatt taccttattg	300
60	tgataatact gttttttttt gtttttttc taatgttgc tttgtgaaag ccaaatccct	360
	ttcagcagca ttgagctaaa aaaaaaaaaaa agtgcgtt tagggctggg cacgggtggct	420

	catgcctata atctcagttac ttcggggaggc cgagggcaggc ggatcacaag gtcaggagtt	480
	cgagaccgc ctggccaata tggtaaaatcc acgtntctac taaaaataca aaaatttagct	540
	gggcatggtg gtgggtgcct atagtcccag ctatgcggga ggctgaggca ggaaaaaaccg	600
5	cttgaacctt ggangcggaa attcccagtt gagccaagat cgcgccactg cactcccagc	660
	ctgggtgaca ganccanact cttgtctcca acaaccagca ac	702
	<210> 43	
10	<211> 642	
	<212> DNA	
	<213> Homo sapiens	
	<220>	
15	<221> SITE	
	<222> (593)	
	<223> n equals a,t,g, or c	
	<400> 43	
20	aattcggcac gaggcgccgg gtcgactgac ggttaacgggg cagagaggct gttcgcagag	60
	ctgcggaaaga tgaatgccag aggacttggg tctgagctaa aggacagtat tccagttact	120
	gaactttcag caagtggacc ttttggaaagt catgatcttc ttccggaaagg ttttttttgt	180
	gtgaaaaatg aacttttgcc tagtcatccc cttgaattat cagaaaaaaaaa tttccagctc	240
	aaccaagata aaatgaattt ttccacactg agaaaacattc aggttctatt tgctccgcta	300
	aaattacaga tggaattcaa ggcagtgcag caggttcagc gtcttccatt tctttcaagc	360
25	tcaaatatccc cactggatgt ttgggggtt aatgatgaga ctattggatt tgaggatatt	420
	cttaatgatc catcacaaag cgaagtcatg ggagagccac acttggatgtt ggaatataaa	480
	cttggtttac tgtaatagtg tgctgttcat ggaaaccgag ggctgcatct tgtttatagt	540
	catctttgta ctgttaatttg atgtacacaa cattaaaagt actgacacact gaaaaaaa	600
	aaaaaaaaaaa aaaaaaaaaaa aaagcggccg ccgaatthaag cc	642
30	 	
	<210> 44	
	<211> 1219	
	<212> DNA	
35	<213> Homo sapiens	
	<220>	
	<221> SITE	
40	<222> (25)	
	<223> n equals a,t,g, or c	
	<220>	
	<221> SITE	
45	<222> (26)	
	<223> n equals a,t,g, or c	
	<400> 44	
50	aattccccgg tcgacccacg cgtcnntaa aatccccaaa ctgacaggta aatgttagccc	60
	tcagagctca gccaaggca gaatctaaat cacactatcc tcgagatcat gtataaaaaag	120
	aaaaaaaaaaaaga agtcatgctg tggcccaat tataattttt ttccaaagact ttgtcaca	180
	actgtctata ttagacatcc tggagggacc agggaaatgtt agacacccaa tcctccakct	240
	tttcagttgtc cctgtatgtca cctcatgatt tgctgttactc ttccatctc ctgcgc	300
	gacagtgggt tctgtgtcca cctttgtgtt ttgcgaggcc gagcccaaggc atctgctc	360
55	ctgcccacccgc tgaccagaga aggtgttca ggagctctgc cttagacacg gtgttacagt	420
	atgaacacac agcagaggca ccctcgatgt ttggaaatgt tgcttctga aagggcacac	480
	ttttaaaggaa aaaaaaaaaaaa atgtaaaaact atactgaccc gttttcagtt taaaagggtc	540
	gtgagaaaact ggctgggtcca atgggattta cagcaacatt ttccattgtt gaagtgggtt	600
	agcagcttc ttctgtcagc tgaatgtttaa ggatggggaa aaagaatgcc tttaaatgtt	660
60	ctcttaatcg tatggaaatgt tgagctatgt ttggaaatgt ccctgggtttt aatccatata	720
	caaagacggt acataatcc acaggtttaa atgtacataa aaatataatgtt tggattttt	780

	tgctctactg tttacattgc agattgctat aatttcaagg agtgagatta taaaataaaat	840
	gatgcacttt aggatgttc ctattttga aatctgaaca tgaatcattc acatgaccaa	900
	aaattgtgtt ttttaaaaa tacatgtcta gtctgtcctt taatagctct cttaaaataag	960
5	ctatgatatt aatcagatca ttaccagttt gcttttaaag cacatgggtt taagactatg	1020
	ttttggaaa aatacgctac agaattttt ttaagctac aaataaatga gatgctacta	1080
	attgttttgg aatctgttgt ttctgccaaa ggtaaaattaa ctaaagattt attcaggaat	1140
	cccccatttga atttgtatga ttcaataaaa gaaaacacca agtaagttt ataaaataaa	1200
	aaaaaaaaaaa aaaactcgaa	1219
10	<210> 45	
	<211> 437	
	<212> DNA	
	<213> Homo sapiens	
15	<220>	
	<221> SITE	
	<222> (422)	
	<223> n equals a,t,g, or c	
20	<220>	
	<221> SITE	
	<222> (423)	
	<223> n equals a,t,g, or c	
25	<220>	
	<221> SITE	
	<222> (427)	
	<223> n equals a,t,g, or c	
30	<220>	
	<221> SITE	
	<222> (437)	
	<223> n equals a,t,g, or c	
35	<400> 45	
	gaattcggca cgagggcggc accagggagc ctggggcgccc ggggctccgc cgcgacccca	60
	tccggtagac cacagaagct ccgggacctt tcggcacct ctggacagcc caggatgctg	120
	ttggccaccc tcctcctctt cctcttggaa ggctctctgg cccatccaga ccggattatt	180
40	tttccaaatc atgcttgcgaa ggaccccca gcagtgtct tagaagtgcgaa gggcacctta	240
	cagaggcccc ttgtccggaa cagccgacc tcccctgcca actgcacctg gctcacaaaa	300
	agagtgcacaa aatgcttctt attccatagc tacggcattt ctcagtaagt tgaggtaaaa	360
	aataaaaggaa tcatacatct caaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa	420
	annaanaaaaa aaaaaan	437
45	<210> 46	
	<211> 533	
	<212> DNA	
50	<213> Homo sapiens	
	<220>	
	<221> SITE	
	<222> (305)	
55	<223> n equals a,t,g, or c	
	<400> 46	
	gaattcggca cgaggaccct atcttacaaa aaagaagaag aagaagaaaa ccatgacagg	60
	tgtcttttaag ctgccttgc tggttgggt tcatgaagca tctgtggag gttgccata	120
60	tgtaaaattttt gttgagttt aagaaatgtt aacgttatat ggtattttttt taatgggtt	180

	ttaaaaataa ttttctcat tcaaattctg aattagaagt tgtttgat aaatattgaa	240
	aattgtttag gggagaattt attcaagtt taatcattt ctttatctat gtataactta	300
	gctantagtt actgaagtgc tcagaattttt ttttagatc ttaactagag tctaaaggaa	360
5	ttactaaaag ctatgtttca aataatatgt aagagtaaag tcctgatgtt aaagattttag	420
	catactgaat taacttagtt gactgatgct gtacttacat gggcttcata tttcttgtgg	480
	ccaagatagc atcaacagaa aaaaaamaaa aaaaaactcg agggggggcc cgg	533
10	<210> 47	
	<211> 1849	
	<212> DNA	
	<213> Homo sapiens	
15	<220>	
	<221> SITE	
	<222> (222)	
	<223> n equals a,t,g, or c	
20	<220>	
	<221> SITE	
	<222> (1300)	
	<223> n equals a,t,g, or c	
25	<400> 47	
	gtttttaaaaa aattaaacaa ggctttgtgt tcctagaaga gcttcatttc agtgaatctg	60
	gtgaccccca tctgttgct gtcataaccgc gacacggact tattttgtc attagcaagg	120
	gggaaaaggc caaaggacaa gggcctttc tcccattgtt tttctgtgg gcagaaggcc	180
	tgaggaatgg gccccagccc gtgggggctg ctgggtcacc ancaggggt agggtgcaat	240
	ctgggtgtgt ttccagcagt gagacgggtgt tattgtgaag gtggcattca tctgcggacc	300
30	aaaacccagc catcggggaa gggtcaggcc ttctgtggaa cttggAACGT gccaggacca	360
	cctgaaaaag ccagggtgcg ttgatcattc tcagatcattt gattggcctc cacttggta	420
	tgtgaattttt tcatgtccca gaagacccaa aagtgtctg gttctgatgt gaggattttta	480
	ttcgtttct gtttccgaaa cacttagcaa agaaggtcac agtgcgtgg agtgcggcga	540
	cccatctttt aagatagccca gtgtccctgg atgaggtgtt gattttccgtt cccaggact	600
35	ctgtgaagtt tagagtacag ttgttgggg tccaaaaggac accatctcta ccccacccaa	660
	ataaaaaatgc actcatctt ttagaacatc tgctgtcaaa ggccagccctg tcgttagggc	720
	atggcttatg cttgacaaac cagtaacaac tggatgtgg cgatgggtgg atgtgtcgca	780
	agcaatttcac tagacaatct tcacatgaat gtcggtagcc aggtcttc cccaggatg	840
	gcttttagtct ttagtgcattt gaaaccatgtc ggaattttta ggtagaaaacc tgggtctgggaa	900
40	ggccctccggac cccaggctcc atccccggct tccccaggcc gcccccccaa gaaaaaccaa	960
	gcccggatgg cagcttagcac ctttcataatc catccccgtt ctcagggga caacaccatg	1020
	gacagccgtt ttcagagccct ccagcatttgc cacaccacta ctcacccctct ctgctgtgg	1080
	catgttggta gagtcatccc tgaatcaag aaatggccctg tggaaatgttta ttgttcaacg	1140
	ttgtttacag ctctaaaaac atggtgagga atgcctaagt cttatgtgacc aaacgtgacc	1200
45	ttgaaagcag acatagcatg acagacccctc ctagatgtt tggctgggtt cacagtgacc	1260
	gagagtcagg tccagcacac acctggggaa gggatgtgn cccaaagggg accaaaaagg	1320
	ccggacgttta cagggtgaaa ccctctgacc cctcgcgaca ccgttaggact tgactttgt	1380
	tttagtcttc taagaaaatag atcatggagc caagtgaagt gcacttgc taaatgttaagg	1440
	gtctgttttgc ttcttgttgc tttctgttt ttttacctt tttccggca tttaaaaaaaaa	1500
50	aaaaaaaaaa aagcttatgt ttcttgtcaat atgcagaaat gttccctccg ccactctactg	1560
	aagtttgca ttctggcttgc tgcagttttt attgtctgtc tcagacgtac agccagacat	1620
	gttctctatt ggcatatccc cgttctgtt cagatgacag cgaccgcctt ttcatcccc	1680
	ccggccacccctg tactcaccctt cacgcctttt gaagaaaaaa aaaaaatca ctttgcgtgt	1740
	tgttagctcat ttgttcaag agagaatcaa cagatcatat tcagtgctt gaataaaatttgc	1800
55	ctctattttgc atattagaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa	1849
60	<210> 48	
	<211> 926	
	<212> DNA	

&lt;213&gt; Homo sapiens

&lt;400&gt; 48

5	ctcaaccaca actagaattt gcacaatata agcttgaac gaaattcaaa agtggtttaa	60
	atgggagcat cttggcttag aggaaaagaac ccctccgatg cctaataaaag ttctctagcc	120
	cacatcttct ggaagcattg aaatccttag caccagcggg tattgcagat gctccacttt	180
	ctccactgct cacttgata cccacaaga gaatgaatta tttaaaatt agagataaaat	240
	aagacgtcg tggtttctta agcacagctc ctcccttctt atattgcaca tgcaacttcag	300
10	ttcatggcta gctgtatagc ttccgtctgt aaacttgtat ttcaagaat ccttggattt	360
	gaatttttag aaatgctcac ataattgttg ggactgattt attcctccac gatatgcctc	420
	ctctctctga tatcctgta actgttagccg ttgtggcatt tgagatgaca ggacatatat	480
	atatatggcc ccacacttga ccttggatgtc ctgaatgctc taaaatcaag catatggcac	540
	agcgctcaag acttttgggt ttgtgtcctt ttttctatgg ctgtctcttc tcaattctgg	600
15	agaggctctgg ttccagttggc tggtttcyar ggattgattt ttaagctctg gatcacagag	660
	agaagcaaca aggaactata ctcaactcaa aacttttag gagaatcatg aaatttgtct	720
	attcaaaaggta tggagtttag tccatwmtgt tattgttgc aagggttgca tattttgtga	780
	gtcagtataaaaatagt ttcttattgt aaatatgata ctctctataa tctattttat	840
	catgtgtata acattcaaac tgacaatattt attgactttat gaataaaagggt gtcaaaaaaac	900
20	aaaaaaaaaaaa aaaaaaaaaaaa ctcgta	926

&lt;210&gt; 49

&lt;211&gt; 1593

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 49

30	gcggacgcgt gggctgtgct ccctgcagtc aggactctgg gaccgcagg gctccggac	60
	cctgactctg cagccgaacc ggcacgggtt cgtggggacc caggcttgc aagtgcgg	120
	cattttctct ttctttctcc ctcttgcagtc cttctgagat gatggctctg ggcgcagcgg	180
	gagctaccgc ggtctttgtc gcgatggtag cggcggtctc cggcgccac cctctgtgg	240
	gagtgagcgc caccttgaac tcgggtctca attccaacgc tatcaagaac ctggccccac	300
	cgctggccgg cgctgcgggg caccaggct ctgcagtcag cgccgcgcgg ggaatccgt	360
35	acccggccgg gaataagtagc cagaccattt acaactacca gccgtaccccg tgcgcagagg	420
	acgaggagtg cggcactgtgat gactgttgcg ctatcccac ccgcggagg gacgcaggcg	480
	tgcaaatctg tctcgcctgc aggaagcgcc gaaaacgctg catgkctcam gctatgtgct	540
	gccccggaa ttactgcaaa aatgaaatat gtgtgtctc tgcataaaaat cattccgag	600
	gagaaaattga ggaaaccatc actgaaagct ttggtaatga tcatagcacc ttggatgggt	660
40	attccagaag aaccacccctt tcttcaaaaa ttgtatcacac caaaggacaa gaagggttctg	720
	tttgtctccg gtcatcagac tgcgttgcag gattgtgtt tgctagacat ttctggtcca	780
	agatctgtaa acctgtcccg aaagaaggta aagtgtgtac caagcatagg agaaaaggct	840
	ctcatggact agaaatattt cagcgttgc tttgtggaga aggtctgtct tgccggatac	900
	agaaagatca ccatcaagcc agtaatttctt ctaggcttca cacttgcag agacactaaa	960
45	ccagctatcc aaatgcagtg aactccctttt atataataga tgctatgaaa accttttatg	1020
	accttcatca actcaatccct aaggatatac aagttctgtg gtttcagttt agcattccaa	1080
	taacacccttc caaaaacctg gagtgaaga gctttgtttt tttatggAAC tccctgtga	1140
	ttgcagtaaa ttactgtatt tgcataatttctc agtgcggcac ttacctgtaa atgcaatgaa	1200
	acttttaattt atttttctaa aggtgtcgca ctgccttattt ttcccttctgt tatgtaaaatt	1260
	tttgatgttta tcttgcactga caaatattctt atattgaact gaagttaaatc	1320
50	atttcagttt atagttctta aaagcataac ctttacccc atttaatttctt aggtcttaga	1380
	acgcaaggat ctcttggaaat gacaaatgtat aggtacctaa atgttaacat gaaaataacta	1440
	gcttatttttc tgaaaatgtac tatcttaatg cttaaatttattt atttccctttt aggctgtgat	1500
	agtttttgaa ataaaaatttta acattnaata tcaaaaaaaaaaaaaaaa aaaaaaaaaaaa	1560
55	ctcgaggtcg acggatcga taagcttgcattt atc	1593

&lt;210&gt; 50

&lt;211&gt; 978

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

<400> 50  
 gaattcggca cgagatgagt ttggccacgt gatgcaccag ctctgctccc aggtgggtgc  
 5 gggccggggc aggggcaggg gcaggggcag gggcaggggc tgcctgtggt cagcgaggcc  
 caagcctggg gcttcggctt ccggctctct ctgcaccctgt cccgtggctc cttcatccaa  
 tgccacccaa agatggtgac tccctgtcar gcccgtgtcc tggggctgccc ccagcaaaac  
 accacagacc agggcttaca caaggtgcgt gtatccctc atggctctag aggctggagt  
 10 cggaggtcac agtgtcagca gggttggctc cctcgargtc cctccttggc ttgtggccgc  
 caacaacttc ccgcacatctca tgtggctgtc cttctgtgt ggtccccat tygtcttctt  
 acrrggacccc agtctgccgg atccgggccc gccaacaac ctcacttgac ctatgtac  
 15 ccttagacat ctgtctctaa gtagtcacat ctgggattac ggcgtgagcc atgttcccgc  
 ggaatttttt ttttatagta ttggataaaag tttgggtttt ttacagagga gaagcaatgg  
 gtcttagctc ttctctatt atgttatcat cttccctttt ttgtacaata tttgttttac  
 ctgaaaaggaa gtttcttatt cgttgggttggac aaagtccaaag tctgtggaaac  
 20 taaaaacctt gaagtcgtt cataggactc tggacaatct cacaccttag ctatcccag  
 ggaaccccgag gggcaactg acattgtcc aagatgtct cctgtatgt tagtggatata  
 aaaggaaaagg ccctgcacag gtggctgttt cttgtctgtt atgtcagagg aacagtccctg  
 ttcagaaaagg ggcttcttctg agcagaaaatg gctaataaac tttgtgtga tctggaaaaaa  
 aaaaaaaaaaa aaactcgaa  
 <210> 51  
 <211> 433  
 <212> DNA  
 25 <213> Homo sapiens  
 <220>  
 <221> SITE  
 <222> (424)  
 30 <223> n equals a,t,g, or c  
 <220>  
 <221> SITE  
 <222> (430)  
 35 <223> n equals a,t,g, or c  
 <220>  
 <221> SITE  
 <222> (431)  
 40 <223> n equals a,t,g, or c  
 <400> 51  
 cggccgcctt agaacttagtg gatccccgg gctgcaggaa ttccggcacga ggcgggaagg  
 45 cttattccaa ggttaagaggg gctgtgtgaa gggcagtgg gatggaatgg ggggtggcat  
 gggacaggca caagggaagc ctccagcccc ttttctgccaa caagcaagag gcaactcagcc  
 ctacctgaga tgtgttattt tttagaaataa tctttatttgat tggtttttgc actcaatata  
 aaggcagcat atgggtgttg caatataaaat ggtacagaag tccacagagc aaaagggcca  
 50 gtttctgtcc ctttccctt ctccaggccctt ctttctggaa cccattattt ggatagatta  
 agaccttcc agacattgtaa aaaaaaaaaaa aaaaaaaactc ggggggggsc ccggaaacca  
 atngcccn naa  
 <210> 52  
 <211> 861  
 55 <212> DNA  
 <213> Homo sapiens  
 <400> 52  
 gaattcggca cgaggctgag tcaacttgat atccaagctt tttacttcaa ttatctggca  
 60 agattacata gactgtcaaa gtttgtgaaa gtttagcaag aaaactgtct tactcacaga

	accacaggac taactgactg aaccacactc caccattgc ccctattcc aggcgttatg	180
	gtcacccctgt agtttctaat ctgtatagat gtgttagagca tgccttcc ctcttccttt	240
	ccccctccctg ttttcccttc ctcttgcct ttcttaatgt ctgttctat tggcttcttg	300
5	atcttggtct ttaatgttca tccttaagct tgcttctctc ttcaagactac tgattcagcc	360
	tcttgcattt tctttcaact tgggcaaaaaa aaacaggcaa cattttctc ctccactacc	420
	tcatcatcat ccaattttat ccttagttt atattaccac aactctccca aacgtcccaa	480
	gtcttattt aagtctaaca acttagcttc gaacctaat ccaagcatct gacaacacac	540
	tggaaatgtgc aagcaaggt cccwatggcc gggtgcagtg gctcatgcct gtaatcccag	600
10	cactttggga ggccaagggtt ggatcacctg aggtcgggag ttcgggacca gcctggccag	660
	tatggtgaag ccatgtctmw actaaaaata caaaatttagc cggacattgt ggtgcacgtc	720
	tgtcatccca gcaaggcagg cgaatcgctt gaacccggga ggcggagggtt gcggtgagcc	780
	gggatcgtgc cattgcactc cagcctggc aacagagcga gactccgcct cattaaaaaaaaa	840
	aaaaaaaaaaa aaaactcgta g	861
15	<210> 53	
	<211> 510	
	<212> DNA	
	<213> Homo sapiens	
20	<220>	
	<221> SITE	
	<222> (380)	
	<223> n equals a,t,g, or c	
25	<220>	
	<221> SITE	
	<222> (396)	
	<223> n equals a,t,g, or c	
30	<400> 53	
	gatccccccgg gctgcaggaa ttccgcacga gtggaaaccg cctccaccaa caccgggtt	60
	tgcctacacc acccccccatt tacttagtat gtttattttt tgggtgtctc ttgccttcct	120
	cccacccctt atttcccttc agagctgtga atggcaggt ctgtctctgg tttggcatca	180
35	ctgagttttt cccatgcatt ggccccaggg ctgcttaggat gtgagacaaa tctccctaca	240
	atgggcttgc tcccatgtc tgtagcttt aatagatgct ggcatgtcgg aggttaccca	300
	tgagtcaaaa tccgctctcc atgcttactc ttgacaccccc attgaagcca ctcatgtgt	360
	gtgcgtctgg gtgtgaagtn ccagctccgt gtggtnccctg tgcttgtact gyccctgctt	420
40	tgcagttctt ttgcacttac tcatcgagtg ctgtttgaa atgctgacat tatataaacg	480
	taaaaagaaaa aaaaaaaaaaaa aaaactcgta	510
45	<210> 54	
	<211> 309	
	<212> DNA	
	<213> Homo sapiens	
50	<220>	
	<221> SITE	
	<222> (301)	
	<223> n equals a,t,g, or c	
55	<220>	
	<221> SITE	
	<222> (305)	
	<223> n equals a,t,g, or c	
60	<400> 54	
	gaattccccgg ggataaaattt cattcccaa agatgagtag gtatgaaaaa taatacttag	60
	aagagattgt tcttgcgggg agaactgctt ctacaggatc tagctttgat tttgtatctt	120

	tcaatcttt taaaatcaac tttaacgaat tttaaacatat tttaagtgtta caagtaataaa	180
	gtttgacaat tggatgtac ttcttaccaca ataaaatata gaacattttt atcattctat	240
	aaaaaaaaaaaa aaaaaaaaaac tcgagggggg gccccgtacc caattcgccc tatagttagt	300
	ngtancgtc	309
5		
	<210> 55	
	<211> 1585	
	<212> DNA	
10	<213> Homo sapiens	
	<400> 55	
	ggaatttctt aaatatgttc atgtataata cttgatcaaa atattttgg gttttttgtt	60
	ttgttttaat gggtagaaa atgtttacaa tcttggtctt atatgtcac caatggaaa	120
15	gtacttcca gttttatatac aatatgagct gactttaact gagttgtttt ggatagggaa	180
	gaagcagtcc ctctacagta tacaactact gcttgcacg tggatcaaaa taatcatgtt	240
	ttatgaaaat atctccctta agcagtgtta aggttgggtt gcagtgtgtt agtggcacat	300
	tgaactggaa gttttcttga aagctgcttc atctattaag aagcaatttt caaattgttag	360
	cgaattatata tttttttttt tttaaagaaa cagtcgttata atgtgtatgt ttctttaaaat	420
20	aactaaaatg tkccctttaa tttttttttt aatggagttt tttttttttt cttttttttt	480
	agtaaagaat ctcttagagg gaaacattt tgctttttt gataatcttc cttgtgcctc	540
	actacatccc taagtgggtt tgactctttt tattaccaca tgctttttt gtatatttca	600
	caaattttact tttaaatatt attttagata cgggttaaca tgtgtcaattt agaataat	660
	tataacaggt catgaaaaac ataactttt taggattca caatattttt wctccacata	720
25	atgagagaat gaatgagcct ttggagatac tgatataagg caattttttt ttgtcaatgtt	780
	gaatgtgttt tttagtttga ttctttttttt ttcccccattt agggcactac ctgccatatc	840
	atctttgtt acttttttactt gtaaagcgcac taatatttac actatgcccattttttttt	900
	attatagttt gtaaattatgtt aagatccttgg aatttttttactt agatctacaa ctactaatgt	960
30	aacagacaag ggcacatcttgg tttttttttt ctgagcatgg cagtttttacc ataaaaaagta	1020
	ctctttttt ctaatttttca gttttttttt aataacattt ctgttactt gacataactaa	1080
	tagtcaacttca agcagtacca tttttttttt ttttgcata tttttttttt tttttaatttta	1140
	atgtatgttgg tttttttttt gttttttttt gttttttttt gttttttttt gttttttttt	1200
	tcaaaagatgc ataaacagcttca ttatcttaggg gaccmccaaa tttttttttt tttttttttt	1260
35	aattttttttt aatattttttt tttttttttt tttttttttt tttttttttt tttttttttt	1320
	tgtatattttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	1380
	aaatcttgct tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	1440
	ttttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	1500
	tgtttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	1560
	aaaaaaaatttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	1585
40		
	<210> 56	
	<211> 874	
	<212> DNA	
45	<213> Homo sapiens	
	<220>	
	<221> SITE	
	<222> (468)	
50	<223> n equals a,t,g, or c	
	<220>	
	<221> SITE	
	<222> (501)	
55	<223> n equals a,t,g, or c	
	<220>	
	<221> SITE	
	<222> (546)	
60	<223> n equals a,t,g, or c	

	<400> 56	
5	aggggaaatct cggtgctgca acgagtgtgg ggccagccgt ggaggctcca ggtttctct ctgccccagc agagccccc aggagcccca acaggaagcc agcgccgcgat ggctgccacc gacttcgtgc aggagatgca cgccgtggc gagaggctgc tgctcaagct gcagagactg ccccaggctg agccccgtggc gatcggtggc ttctcagtca tcataccttt cacagctact gttctgctgt tgctgctgat agcctgcagc tgctgctgca ctcaactgctg ctgcctgag cgagaggcca ggaaggcca ggtgcagccg acaccacat gacggacggg cgatggctga ggagaagctg gagaggagat ggccaatgcc atgacacagg ccattcagctt ggcctgcag 10 cccttacccc tcaagaccag gctccctgg cccagctct ggcccagncc caggtacctg gacactgaca acttgagccc ntaccaagga aacaagggtt ggtatagtg caaacctctc atctgnccag tggacactgg gtgctggga gtcagctgtt tcaaagactg ggtcaactgc ctgggcttct tcgccttaccc gcaccccttta acaaaaacaag gaagtagggg tcccatacc ttgatggaga acagtccccca cctgtggca attggccctt ggggctctgc tgatacatgc 15 caaagaggag caaggcaatc agaggggctt tggcaatag cttctgcata cgagctcccg ccagagctgt agcatgtcag tattctagtc cagttttgc cagtttccaa gtaaaagctt ttgtgttaaa aaaaaaaaaa aaaaaaaaaact cgta	60 120 180 240 300 360 420 480 540 600 660 720 780 840 874
20	<210> 57 <211> 1169 <212> DNA <213> Homo sapiens	
25	<220> <221> SITE <222> (2) <223> n equals a,t,g, or c	
30	<220> <221> SITE <222> (9) <223> n equals a,t,g, or c	
35	<400> 57 gngccggccnc cttttttttt tttttttata ttttatcaat tttattgaaa tattccaagg atcccaaccc catttaaaaa taaaaattgt aaagcactcc attcaataaa agcacataag tccccctcaa taatttagtat gacaattcac gatacagctt ttactctggg agagtttatt ttacccttta ttccaaaagg cacaaggta tctgaggccct cagatattaa cccactgca 40 tgtaatgac acaccactga ggtcagctc aatgtatttta ttaaagcttta taacacactt ccccaagaat ttatagatttcc ttcttataaa taataattttt aaaaatactg caccttaaga ccaatacagg cttAACAAA gacctgaaat ttctgcaagg gcagtttgc ttcttgcatt aagtacaact ttgaaatgtc tattcccagc aaaagaaaca ctagaccagg ctggccaaa gaaacaaaat aaaacaagtg atttctaaca cgctaaaaga gtacattttc atcagctcca 45 aagaaagcag ccctggcat tcagaaggct cctatgatcc caccaggctg cagtcatttag aaatatatgc ttacaggcc acaggctgt ctggatttttgc ttctgacac cagtgaccagg aaaagccag ttttgcgtgt gagggggtgt ggccccccgt gccttggcc tgctcaccgg ggtgatgga ccccccgggtt gtacacaggct gctgtcacgt ctggactgtt ggcctcttct 50 gcatctggcc tggggcttc tcctgcttc tggcccttag tcacgttatt gtctggctgt ccgggtctgg ctgcactctc atttgtgagg ataaccctt ccttcttctt ttctcccaat acctccagcc ccatcatctt gagataatga agccgttcat tttggccac aaaagttcga atggaggcct ttccccggca tccgcataag acgatgggac actgcagagc gtctggattc 55 gcagaatctg gttcatactt cagcacgtt cttccctttt ccaggtccctt tgcttgactg taggtctcac tgctgagttt tctaaaaaaag ggattttccctt gggtaacacag tatcttaaca tttccatgtt atacagtaat aatttttg	60 120 180 240 300 360 420 480 540 600 660 720 780 840 900 960 1020 1080 1140 1169
60	<210> 58 <211> 1066 <212> DNA	

<213> Homo sapiens

<400> 58

gaattcggca	cgaggcaaatg	ttgaaccaat	tatgttttg	tgggttgtt	cttagctgt	60
gaatcctgaa	tggtttataa	agtgaactag	ctggcttaat	gcagccagcg	ttctgggcag	120
cagaacatat	tcattcttac	tgtaaattct	atttgctgt	tccaaagggtg	atgattttca	180
agcagacatg	ttctatatatgg	tctgtgttt	aggatctgtt	gcccaagccctc	tatcagagct	240
tgcctacctg	gcaaagctgc	ctacccttca	agtggggaaa	tataatccac	tgtttaacaa	300
ggctcacccct	ctccacccctg	tcctaacgac	cttttgtgaa	tgtgctgtga	tattttcttg	360
ctcaatagca	agggtgttagc	tctgcttca	tttaagaaa	gtggaggctg	agggcattgt	420
atcaatactg	ttgaacttcc	aagaagttt	ccttgtaaaa	ttaaaggaaa	gatcttgtta	480
ttgattaacc	attttcttat	gccttgctat	tgacatattc	atgcctttc	tacgtctagt	540
ggctgaaaat	gtttgcattt	gttcatttga	ctaatgggt	gattttgky	ycwatattat	600
tagacctgta	atgttttaaa	atgtatttta	ttaaatttgg	actggatgta	tgkccctctag	660
caatacggagg	tactttctaa	actattaagg	gaggggttgt	aycctcatgt	tgagataaga	720
tgatggctgt	ttaaatttttgc	caattttttt	tggcctgcag	ggatatttttgc	tgtttatgtt	780
tccaaaaaag	gaataaatttgc	gcattcttgc	gccaaaagtt	gtttttccctg	tcaattgtct	840
aataagttatg	cagtagactg	taatggcaac	atacatggtt	gctttataaaa	aacagtttcc	900
tcagttatgag	aaattttaca	aagaacagtg	aaaaaaacttt	gtgtttttaa	ctcttgggtc	960
tcccttatttt	taaaaaatttgc	tatgggtat	acaatttatta	tgtgtcaatt	aaaactaaaa	1020
taaaaactttt	aaaaaaaaraaa	aaaaaaaaaa	aaaaaaaaaa	ctcgta		1066

<210> 59

25 <211> 772

<212> DNA

<213> Homo sapiens

<400> 59

30	gaattccggca cgagctttcc tgagcctcg tttctccaac ggtgggaggt ggttagaaatt gatatagtac ttaccactga gggtaaaatg agatataacc tgtgtaaaata ctgtacacca cagtcattca atagtggcag cttaaaaaaaaa ttattctacg attacccttg cttcagtgat tcttcttgggt gttatttgaag ggtgagatct cggtggggat cttccagggtg tttccataat cccagcgatc accccccaggga gaacctctct ccttaggctg cttagaggaca tggccatag	60 120 180 240 300
35	gaccagatag gaggggggggg cagcggtggg aatgcgtttt cagagetacc tttggccaag ccgtatccct gtggggacct attgcattgc tgctgaagt ctgtcccat cagccctggc ttcggtgtggc cctgtctggc aagggggtgc tcctacaaaag tcatggcagc ctggtgccaa aaccatcatc ccataggacc tgctgttagct ttggccagaag cctggccaa ggggtggagg ccctggagg tctgaccac cacgtggagg gtggggaaatg ccacagagca ggttctctag	360 420 480 540 600
40	aagggatttg tcagaagcta aactgggtg cccccctggc tcaggcctgc acagtttctc cctgaccacc cagctgggat ggatatacgag acaggtgtca tggcagaa agcctgcct aagaggccct actgggtgtt tcctttatta aaaaaaaaaaaa aaaaaaaaaactc ga	660 720 772

45 <210> 60

<211> 1198

-213- DNA

<212> DNA

50 <330>

<221> SITE

<222> 118

<223> n emials at g or e

55

<221> SITE

<221> 311E

<223> n equals a.t.g. or c

60 <400> 60

	tcgacccacg cgtccgattt aattcttatt ccccacagtt taggtatttt tcatttagtac	60
	atcaatttga cacactaat gcaagactat taaggaagaa cgattaaata ttattttatt	120
	ttgtgaagag ttggcagcac attacatctc aagaacttgc agagagagga aggttagatgg	180
5	acaatcctaa attgttaagat gttacaaaaa acagtgaagt aagagtactc ctgaagacta	240
	aaatagagag gctggggttt gagccatttt actgagtagc ttagctggaa cctgatatac	300
	gaagttagcct ttaacaaaaa gcctcttgc aattgtatgg tactaacaac tagagtagt	360
	aagtgttaagt taaaaccaag ttgcagtggg aaatcaaagg tgaggtagct tatttggaaac	420
	cagcaaatga gacagggttg acagtttaa aatctttct aacaaagaaa ctgcacggta	480
10	gcaaggacta gcggttctca aagcccttct tttcagtgt tctcattcac cttggcaccc	540
	aagtatgtt aacaggccat gcattaaaaa taaatacataa aatataaaaag ccgcttaaag	600
	ggaacttaca aactgacaat ctctctctg tatttgggtt catagtggct gggagttaa	660
	ttatatgcac aaaagtttagg agccacttgtt ttcgcacag actgttaggag caagatgagg	720
	agatggcag gttttggtaa gagccccag ttctggtgga caggcatact tggcatttg	780
15	ggcggccat tgctgggagg accacgttcc gggaggcgat tgacttttg tttgttaattt	840
	cccttaaac aagaagagat ggctcacattt tccatataat atctcaatga atgtactgta	900
	ttactgtttt aaaaatttga taaaataata atgaatttggt ctccctttgt tatctggtcc	960
	ttgtttaattt tggtaaaggg tttttgtata caaaagtttta catttttatg tatatttttc	1020
	ttgtgtaaaaa actgtatgtt aatgtgtatg aacactgtt tttttttttt gtatataatgt	1080
20	tggatgtttt tttttttttt gatgtattaa taaatcttgc tggtaagttaa	1140
	aaaaaaaaaaaaaaa aaaaaaaaaactc gaggggggggc ccggtaccca ataaccctnt natgatct	1198

	<210> 61	
	<211> 558	
25	<212> DNA	
	<213> Homo sapiens	
	<400> 61	
	ctgcaggaaat tcagcacgag ytggcatgtg acaacccagg gctgcctgaa aatggatacc	60
30	aaatcctgta caagcgactc tacctgcag gagagtcctt cacatcatg tgctacgaa	120
	gctttgagct catgggtgaa gtgacccatcc gctgcacccctt gggacagcca tcccaactgg	180
	acggggccctt gcccgtgtt aaagttagcag aacgcggcagc agagacgtcg ctggaaagggg	240
	ggaacatggc cctggctatc ttcatcccg tccatcatcat ctccctactg ctgggaggag	300
35	cctacatatta catcacaaga tgcgttactt attcacacccctt ccgcctgcctt ctgtatgtact	360
	cccaccccta cagccagatc accgtggaaa ccgagtttga caacccatttacgagacag	420
	ggaaaccag agagtagtag gtttctatctt aaagagagct acacttgaga aggggacttg	480
	tgaactcaac cacaatctcc tcgaggggg gccggtaccc aattcgscctt atagttagtc	540
	gtattacaat taatggc	558
40	<210> 62	
	<211> 616	
	<212> DNA	
45	<213> Homo sapiens	
	<400> 62	
	gaattcggca cgagtcttga cagccctggc accaagggtt tggaaaaagg ttctatttgg	60
	gtggagattt atgggtggaa aaaggagaga ggggagttgg acctgatacc aaagagatgt	120
	tttcagccat caaccagctg caaaaacaaga tggcttccctt ttcctacat attctccaa	180
50	gcatcataaa tactcggtct gctcccaac ccacatcttgc caggatgcag ccagacaa	240
	agccccactc cactctgaaa ccagtcatttcc tagggatgtat gatcttttct tagttccct	300
	gttggagggtt ggttgggtt ggctgatcgc tggctgggtt actcctgcac tggctggcg	360
	ttggctgcattt ggttaaagctg tttttttttt catccctgttgg ggtttttttt ggtttttttt	420
55	gcatattttcc tccagagacac tggcagacac aaagggtcaa cagaaacctt caaggttttt	480
	tcatgccttac ttttgcactt agcacattgtt catttcagcc tcatgttattt gaccaaaagca	540
	agtcaacttga ccaaatttcaa agccacaaaaa ctcgtgccga attcgatatac aagcttatacg	600
	ataccgtcga cctcga	616
60	<210> 63	

<211> 811  
 <212> DNA  
 <213> Homo sapiens

5 <400> 63  
 gaattcggca cgaggagctt ccatttttc tgatgtgagt ggtgtcagga atgactatgg 60  
 tgggtgttagt ggcaagtggcg atgggtttct ggaggctgaa aggttaaagt cccaatgcag 120  
 aagtgtatgc agggctagtg ggtggcggta gcaggtcag taaaagtcaagg ttcatgtct 180  
 10 tcaatggtga ctcccccttc gtgttagtcc tacagcatca ttccagactt tggttttgtt 240  
 gcttagctcc aaggcctttc ctccctgtgt cctgtcaggt tgggtccact atgatggac 300  
 aagaccctgt catctatgtat gatgtatgc accttgctaa ttatttttct gtttaagcta 360  
 gccatagtgaa atccgttat ttgtgcctaa gagctttac tgacaaagaa cgtgttaccg 420  
 gaagtgggat gctacaagta acaacactaa aagtagaaatt gactaagtgc agcaggcagg 480  
 cctttgagca aggaggggac acacattaca ggctggaaag ctggtgactc ttgtaatgca 540  
 15 gtggcaaaat ttgcttcaa ctactatata caataacttga agatgcacac tgcaagctga 600  
 gtgaggctgt gataagaggg gaaatagtgg ggagcattca gaatgttgg ttacattgt 660  
 gacttcttgc ttttcagca gtcttgatag agcagctata cccacaccag agtcctccag 720  
 ctgacaagag aaggtaaggag agaaaactgct ttgccaggag gggccctctg ctgcagctgg 780  
 20 aggtccaagt tgaccgagag cccaaattttt g 811

<210> 64  
 <211> 993  
 <212> DNA  
 <213> Homo sapiens

25 <220>  
 <221> SITE  
 <222> (370)  
 30 <223> n equals a,t,g, or c

<400> 64  
 ggcacgagcc caaagtgtcg ggattacagg gagttgtatga aagtggagat gtttttagag 60  
 ctacccatgc agcattcaga ttttcttcta ttctgtgtct gctggaaagc catggatcc 120  
 35 aaaaagtctc catcacattt ttscggcagag gttagggggta ttatscccag ttttggatgt 180  
 ttgaatgtca ccctttaaag gagcctcaca tggaaacacac gttgagatttca caactctgt 240  
 gacaaaggcat cgaaggcagaa aatgagctg aaaacgcattt cctttccacg gattccctca 300  
 taaaaataga tcattttgtt aagccccgaa gacaagctgt gtcagargt tctgtcgca 360  
 40 tacctgacan gcagcttgat gtgactgtc gtggagtttta tgccccagag gatgtgtaca 420  
 gtttcttgc gactagtgtg ggggaatcac ggacactaa agtcaatctg cggaaataattt 480  
 ctttttattac acactcaactg aagttttga gttccagaga gccattcttgc tcaaaacattt 540  
 ccaagtactc tttgagagcc cagcattaaac atcaacatgc ccgtgcagtt caaaccgaaag 600  
 45 tccccgagggc aaatttggaaat ctttgcattgt cattcaaaaca gatgaaggca agagtattgc 660  
 tattcgacta attggtgaag ctcttggaaa aaatttaacta gaatacattt ttgtgtaaag 720  
 taaaattatcat aagttgtatt ttgttaactt tatctttcta cactacaattt atgcyyttgt 780  
 atatatatattt tttatgtatgg atatctataa ttgttagattt tgtttttaca agctaataact 840  
 50 gaagactcga ctgaaatattt atgtatctgtt cccatgtat tttacttac ttttacaggt 900  
 gagaagagag ttctgtgtt gcattgatta tgatattctg aataaaatatg gaatatattt 960  
 taaaaaaaaaaaaaaa aaaaaaaaaaaaaaaa att 993

<210> 65  
 <211> 689  
 <212> DNA  
 <213> Homo sapiens

55 <400> 65  
 gaattcggca cgagcttaagg tgggggggtc acttaaggctt cggaaactctgtt gcctcaagca 60  
 atcccttgc ctttcccttc caaagctatg aaatttgcaga caggagccac catgcctggc 120  
 60 tggttttgg gggccatggc aagtgcaggc ttgtcagagg aattggagaa gcagggttata 180

	gttagaaaaa cctctccact tcttgtgtt catgccaggt agtgtttgta acttcagaac	240
	ccgccttac cttacctacc taccatgtt tgctcatttc acctactgtc ccctgctgta	300
	tagggagtgc cttgaggca gagatcatgt tagttttgtt cccttctg tacagagggt	360
5	ggagcccagt acctggcaca gctgaaggag gaatgtgtc ctgctgtctc tgtatccca	420
	gttactcctt gttgacctct agccaagaca aggaacctcc ttatgagatg tcatcttctg	480
	agctcttctt atggaggaa taccacggtg atgattgaat atgaaaagtc ttggcacagt	540
	ggctcacacc tgaatccca acacttggg tggccgaggt gggaggattt cttgaagcca	600
	ggcattgaga ccatccttgg ccaccaaacg agacccatc tctacaaaaa aagaaaaaca	660
	aaaccaaaaaa aaaaaaaaaa aaactcgta	689
10	<210> 66	
	<211> 942	
	<212> DNA	
15	<213> Homo sapiens	
	<400> 66	
	gaattccagg actgctggga cccctgcac ctcctggcca cggagagatc ctgctcccag	60
	ggaccagcgt ctgggtggga cacagttcac tcctctctcc acttcatgtt ctttttcttc	120
20	agcagatggc tcaagtccct tgggtcttc cttgctttctc gacagccgta gcttctgaaa	180
	cctgccattt tgggtctctt gatgctgtat ttcttataattt tcctgactgt gtcttcttagg	240
	aagcattaag tctgaactga cttattaggg aacttcagaa agttaaadaac acaaaaaccct	300
	ttctttgact cctatcttaa ggacatggag atacagttac atatatttac acacaaggat	360
	attcatatgg caaaaacggg gagaaggcac aatttaagag cccaatgggg actgggattt	420
25	tgtatgcattc tgcataatga catgttatga agtcattctg tttttttaaa aacttttttag	480
	tgacatggga aaatacaaag aatgtaaaga attaaaaaaag cagcgtacaa aacmatatat	540
	gtgatccaaat ttgtgttggaa aatattttat ctatataatccattttaaa maccacarga	600
	aaatacacag ttaacagtag ttatctttgg aaggcaggat tataagtgtat cttatgtttc	660
30	ttccctccac tttgttacc gatatcagaa aaaaactctg tctctacgaa aataaaaataa	720
	aatgaaataaa aataaaaattt gctgggtgca gtggctcatg cctgttgcct cagctcctca	780
	ggaggctgag gcggggagaat cacttgggccc cggcaggctcg aggctgcagt gagctaggat	840
	cgtccactg cactctagcc tgggtggcag caagaccttg tctacaaaaa aaaaaaaaaaa	900
	aaaggaattt gatatcaagc ttatcgatac cgtcgacatc ga	942
35	<210> 67	
	<211> 2309	
	<212> DNA	
	<213> Homo sapiens	
40	<220>	
	<221> SITE	
	<222> (13)	
	<223> n equals a,t,g, or c	
45	<220>	
	<221> SITE	
	<222> (652)	
	<223> n equals a,t,g, or c	
50	<220>	
	<221> SITE	
	<222> (677)	
	<223> n equals a,t,g, or c	
55	<400> 67	
	ggtaagagag aangtgtgaa gtcacctgtc tggcttagtc attccaatgg acattgtact	60
	ggcccgaggag gaaagaacca gatgtggttt tccagtcattt caaagcaagt ctctagcaca	120
	aagccccgttc cactgaactg cccttctcca gtgccttc tggatgtggaa tgatgtggaa	180
60	ctcccccttcc ccacggatgt gatccagcat aggttacggc aaatcgaaac agggtaaaaa	240

	caagagggtgg	agcagactacg	tcgacagggtg	cgtgasstca	gatgaggctg	gacatccgtc	300
	actgctgtgc	ccctccagca	gagccccccc	tggactatga	ggatgatttt	acatgttga	360
	aggagtcaga	tggcagtgtat	actgaggatt	ttggctctga	tcacagtcaa	gactgcctt	420
5	cagaagcaag	ctggaaacct	gttgataaga	aagagactga	ggtgactcgc	tgggttccag	480
	accatatggc	atcacactgc	tataactgtg	actgtgaatt	ctggttggcc	aaacgaagac	540
	accattgcag	aaattgtggg	aatgtatttt	gtgctggatg	ctgccacctg	aagctgccc	600
	ttcctgatca	gcaactctat	gaccaggatc	tcgtctgtaa	ctcatgttac	gnaacacatt	660
	caagtctctc	gtgccanggg	aactcatgag	ccaacagctg	aagaaaacca	ttgctacagc	720
10	ttccagttga	atgccgggaa	aaaacctgtc	caatttttagc	aggtttgaag	ggaggatctt	780
	tttcagttgt	agtttggaaag	gttccttggt	gtggctcatg	aaatcacaga	gctcagagat	840
	accatcttga	gaaatcctcc	ttggtatcat	gaaactggag	cagaggaaatt	gcaattttagc	900
	aggaggctct	ctactggta	taccctcacc	ttggggtaat	gggtctaacc	cagaccagg	960
	gtctggaaag	cttaatgtt	agttgggtac	tccagctct	tttccttgg	ggtcacaaga	1020
15	tgtatgattgc	gtagatgtt	cctggtgcaa	agtgcacccaa	acagcaatag	aaaggccat	1080
	gtataaccaa	actccaagtg	ataaccagac	ccatctctcc	tccaccttga	caaaagcaga	1140
	ttatagtata	caaggttagga	attcctgtcc	tatggagat	gaactatac	ctgtacact	1200
	gtgctctgtg	tctgcataaa	ggctcagcc	ttagaggcac	tccttctagt	tgcatttagt	1260
	ctgtctttct	gtggagttt	gtttgaagac	tggctcagca	agtggaggtt	tcaatgtatt	1320
20	tttcagttgg	ctcatcagcc	agcattgggt	aatattttagt	ttagggaaac	agttttaggg	1380
	agtgagacat	ttttgggagc	agaggaaaaac	tctgctgtat	ttcggtcctg	gcaaacattt	1440
	agttatTTT	agctgtgaag	gcagtcgtct	ctgttacaca	gtggcagctc	ttgagttatg	1500
	cactgtgaag	aatgagaagg	gaaaagcaaa	aattatccctt	gtggaaatatc	tgctgattgt	1560
	gccctactct	ttgcacactg	cttttccat	ttgtccttgg	gctaacaacag	gagctacacc	1620
25	tttgatcctct	cctggcatga	aaataaaaaca	aaggtttgc	ttgttgcgt	tccattgccc	1680
	atttccccc	tgttgtctt	cccttggctg	atgcctctc	tgggtcacat	tgcttcttat	1740
	cctgaacact	tgacacctt	agggttagat	ttagcgttgc	gtttttacat	cctagcatat	1800
	gctgttttgt	atgtgagggt	ttcagtagaa	atgctgtct	ctatTTCTGT	gcaactaaca	1860
	atggaaacca	aacagaagag	aataaagcct	tgataccaaa	attgggaaag	aacatgtgtc	1920
30	catttggacc	aaacgttgc	ggtttttaaa	aaattttatt	ttgttttttt	gtttttgttt	1980
	ttgttttttt	tcatcttaat	atgttaccat	ggcacttaac	caaaagatac	agtgtatag	2040
	ccatgtatct	gtctacttag	cgtggctgtt	ttgaggact	gtcccatcag	tgaacaaact	2100
	gcatggcctt	ggagagagac	tctggctct	tggctcagat	gttttcatca	aataactcctt	2160
	tcagagctgt	tgtgggtgt	atgtacatga	tgtggccaaa	aatccaaact	gtgcagttgc	2220
35	gttgtgacaa	acatgcaatg	tgctgtaaaa	attcaataca	gtttaataaa	aatctctata	2280
	ttagtaaaaa	aaaaaaaaaa	aaactcgag				2309

	<210>	68					
	<211>	814					
40	<212>	DNA					
	<213>	Homo sapiens					
	<220>						
	<221>	SITE					
45	<222>	(421)					
	<223>	n equals a,t,g, or c					
	<400>	68					
	tacgagttt	ttttttttt	tttagccata	attacaaaaa	acattagtgc	aggacaccat	60
50	ttaaaaaaac	tatTTAAAT	agtcttcaga	aaaaaaat	taagtattac	agtttaggag	120
	tatattgtact	ttggggcaac	ggattccaa	atTTTACAA	aaggcaat	ccacgcaaca	180
	tatTCAGAT	tcgggttgc	gagaagctgc	agggtttgc	gtgactctat	cacaactgt	240
	ttccgtacgg	aggagccact	gccaactgt	tggacgagaa	tacttaagca	cgtgtttcat	300
	tgctccactg	ccacaggtgg	atatttcagg	ggaattatta	ttaatttcaa	agtttttttta	360
55	aaargytatg	ataagtaat	aaaagtaatg	gtaggaktca	cggcgggaga	gcttacgc	420
	naagtcttcc	tatagccttc	ccccggaa	cccagttcag	gcatcggtca	cccggaaatgt	480
	caccctctga	tcttcccccc	atcccattct	aggaagttaa	agagatccct	cacaggtacc	540
	gtggctctcg	gtgcctctcg	acttccaaaca	gcccgttgc	gcccaggaga	ctcgctccga	600
	cctccaccac	aatggggggc	agtgtgggccc	gcccacccag	aagtgcggcc	gcccaccc	660
60	cccagttcc	gcctgcaccc	agagctcagc	gcaccagccc	ggctcagcca	gacgaaggca	720

	aacgaagaga tgcggatccc tggaggactg gccccaccgt gaacaaaaca ggaagcattc caggaagact gcgggggtgg gctcgccg aatc	780 814
5	<210> 69 <211> 788 <212> DNA <213> Homo sapiens	
10	<220> <221> SITE <222> (370) <223> n equals a,t,g, or c	
15	<400> 69 gaattcggca cgaggcaatt ttcaatgaac cttgaatggt aggaagaatt gaagaagaaa tcagaggatt tttgccttgc agaaggcagc tgctgtgatg gcaggaggt gaaatggaca tggcctggca gaagagtatt atgggggtgg tttgttgta gccatctggc ctgtacaatt tggagaaaca atactttttt ttttcttctc tgcaagctgg gtttcctgtg atttgttct 20 caggctgcac aaaaatagcg tatggctttg ctgtgttattt accttcatct taaaatagct agaacattttt ccctcttctt taaaaaagtt tttaaaatga gggtagact cttgttagaa aaggtagaaan tcttaataac agtactcatg ttgacaaaacc tttctcgatca aaattcctat gtaatcaaga ctcttataa atatgaacaa atgtaatgtt tggaaatcaa ttgttaccct 25 caaggtaaaaa gctgaaaatgg atttataaaag aattattttta aacagcaata atgtttgagg ggtgggggaa gtgagaaaaaa tgaatttta aatcacatgt ttatgactat gaagctagac ttaaaaataa ggtcagttttag ggtatgactc ttataataca aaagtttattt tggatatacaa aggatttataa gctaatgttat ttttaatta tattcaactaa tacttgtaaa agatcattca 30 atttataaaag tttccaaaat aaacctgttt aaagtgtcaa aaaaaaaaaa aaaaaaaaaa aaactcgaa	60 120 180 240 300 360 420 480 540 600 660 720 780 788
35	<210> 70 <211> 791 <212> DNA <213> Homo sapiens	
40	<400> 70 gaattcggca cgagctcaag gctaaaatct tgatctctcc tgaatatgag gaggtgtgtt aggcatgttt tggggattgg attaatagtg taaaaaattt ttttttttca caaaaatagc atgtacccat cacccaaact cagcagctt caagaagctt ttctttttt ctttcttatt ttaaaaaatc cttaacctt atgttagtt tataatcttt taaaaaagta gaaaatcatg taaccttagg atttttagtt ttaatgtaga gttcacaaa tttccatctt tagtaagaca aaagggtcac atattggctg tctccttcaa ctatacttcc ttctgtataa aatatgttta 45 ccatggttgt cattatcgag cacgtaactg catgttagac tctatgtcaa gtgttttaca taatcattta aagctacta aggccctagg agtaattattt atcctcccat caaaaaggta agtgaaatgt taacctgaag tttgactact ttaggtctct gagctagtaa gtacaatagc caggtttcaa accaagatcc ttttaactgc agcacctgtg ctttatctgg tagcgtcatc ttggttcata cattaaaaa agagttatct atgtgccggg tgccctggct catgcctgt 50 atcccagcac tttgggaggc cgaggaggc ggatcaccag gtcaggagtt tgagactgac caataagggt aaatctgtc tctactaaaa aaaaaagggg gggccctac ccaatcgccc aaaaagatcg t	60 120 180 240 300 360 420 480 540 600 660 720 780 791
55	<210> 71 <211> 804 <212> DNA <213> Homo sapiens	
60	<400> 71 gaattcggca cgagcggcac gagcttgaaa tggcgtttc tgatgaacac tcatccatcc	60

	ttcaaggct actctctcat cacagttgt gactcttcca ctttttgaac tggtgtttcc	120
	cattcccagt tcacagagcc ctttctcatt gaactattta tctgagttcc ctctgccgga	180
	acatgagccca tgccTAGAGT agccacctag tagtgagtga cagctctgtg ctggatgcac	240
5	ataaaatggtc tcccttaact gccatgagsc ctaaagaagg tttgctacag ctatTTACA	300
	gatggggaaa actgacagag agatattaaat gaattgccc catgcaaATA tttgctgagt	360
	cttggatTTG catctttatc gtgactccac ggagaccac CCTCTAAGAC cagagccagt	420
	gtccatatca tctttgtct ctgcAGCGTT cagcatggca ctgttggc ttacaaaATC	480
	tgcTCTATGC ttgctgactg ctGAATGAAT gaatGAATGA atAGGTAGTC acaaAGAAATG	540
10	tttagaatgt ttctcagaca ggctgagaaa aaacacaacg aaacattatt tccgtttgga	600
	aagTTTTTTT atTTTGTGT tcAGTACTGA agTAAAACAA aaATCTGAAT aacAGCTGCA	660
	ccgtaaaaaa tgaattacc aatatATGAA ctctaggcat catgcata taatTTTTG	720
	tagataacct ttctctcat tttcTTTCTC attctcttca tcttttctt tttgtttgag	780
	aaaaaaaaaa aaaaaaaaaac tcga	804
15	<210> 72	
	<211> 783	
	<212> DNA	
	<213> Homo sapiens	
20	<400> 72	
	gaattcgca cgagctaaaa cttacaatga catgttggc tttgctctgt aagctccaag	60
	gcattttttt tttcagTTTT aattcaagtG ttctaaaaAG tattttgggt acaaccagaa	120
	ctctctctgc tccttggatt ggagtcaGtg tgaaAGGAAC acagtgggct ctgggggtcaG	180
25	ctagacCTGG atgtggatca cagctcacct ctTCATTGGG aggCCTCAGG caagttattt	240
	gccaacCTCA cctacaaaAG catgatgcta agctcwtttc agtttagttg tggatATCAg	300
	agcatatGTA tacaatgcct gccatAGTGA gtgcCTGGCC ctTGGCAGAC tGTCaaATGG	360
	agctatggag cagcAGCAGG AGTAATATTA ttatctAGAC ctTATCTGTC ctTTTAAACT	420
	cagttcAGAT tccttctcCT ttTTAAATTa CTGCAACCTG atTTTACCTG CCCCTGCCTC	480
30	caagttGCTG tATCAGTTAG CTCCTGAACA attcatttag caattttAAAt tATATATTGc	540
	ttcttgcACAC tgctttgtGA tCTTAAAAAC tCTGCTTCAA atACGTACTT gtttgctttt	600
	cctgAGTGCT gttAATTCCt GCTCTAACGG actAAAGTA tttGAAGGCA ggACTAGGTT	660
	ttatGcatGG cacACAGTCT ggtgcCTTAC atGTAACTAC tcACAAACTT ttttGATCCA	720
	aaATTTAGAA acttcACACG cattcATAAG aaATCAATAA aaaaaaaaaa aaaaaACTCG	780
35	tag	783
	<210> 73	
	<211> 1523	
40	<212> DNA	
	<213> Homo sapiens	
	<220>	
	<221> SITE	
45	<222> (1)	
	<223> n equals a,t,g, or c	
	<220>	
	<221> SITE	
50	<222> (8)	
	<223> n equals a,t,g, or c	
	<220>	
	<221> SITE	
55	<222> (15)	
	<223> n equals a,t,g, or c	
	<400> 73	
60	nggggggncc ccccnTTTT tttttttttt tttttttttt ttcAGTTCTA ctattatTTA	60
	tttttttAAA tattttGAA aaaatataat ttttttacAA tattttcaAC taaaACACTA	120

	ttcacactga	acacgtatgg	cagcttaacc	tacccaaata	tgaagttaa	gaagccaaaa	180	
	ctgttctagc	tttgtaaaaa	gttgctgc	agactctcg	gatggtaac	aaagcaagg	240	
	aaagcaccac	tcaaatcata	atgttacagt	atctttgttc	agctggatta	tgggtggta	300	
	ttggctcatat	gttagactcc	atacaggcat	agctatgatg	cagtgaatcc	cttagaagtt	360	
5	acaattctca	aattacatac	ttcctcagat	gtaacattag	aactcaatat	ttctaacaat	420	
	aacataccag	aaaaggctgg	actggcactc	atctgctgac	taactttag	cctcagtaat	480	
	atgacatact	tgccttaac	aaattatctc	aaattaacta	acagacctc	agaaaatgga	540	
	gattctttt	gatggggaca	taatcaaatt	taagtctgag	aaatatgctt	aacagttga	600	
10	actcaaatta	aatgtactga	ttttaaagtt	tagacattaa	caagtatag	attagcctca	660	
	aaaaaaagaca	atttggtaag	gtttaggtct	tttaatttgg	tgcttgc	caacttgact	720	
	gttgcttctt	tccttgctgt	cttcacatca	agccatgggg	ccaattctat	tttcagtaaa	780	
	tgtttgacag	cttttactt	agtaaacagtc	ttagcacttt	tattaagcat	gcaagactaa	840	
15	aaaaaaactt	ggcaatgcat	aagtgtaaaca	cagtgacaag	agagctttt	caattaagtc	900	
	ttctaatact	gccttcacag	tgtgaaaatt	gtgctacatc	cacccaaaaga	ggggccccgtc	960	
	tactcaaata	tttccgtact	tcaccccccagg	aacaaaactcc	tttgcatttgc	gattcagatt	1020	
	gctcttgacc	acaagatctt	ccagagaaga	gccatcactg	ataacaaggt	cattaaactg	1080	
	gtcttggatt	tggccatag	tttggggag	atctcgagct	ggaataaaacc	attcatgctc	1140	
	ttcttcctct	tccagcattt	cttggaaacaca	gcgttcaata	aattcttctt	cccatataactc	1200	
20	ctcttctatt	tgrctgttga	attcttcttc	attttccatc	cacatgtact	ctgcaaatgg	1260	
	atttgtcatct	tcatgagaat	gaccgttaat	aatcacatct	tcattgtatga	tgcttgggc	1320	
	agtactgtcg	cgacttggat	ctttcatggc	tgatgttgg	tgtcggtttt	aacccaatgc	1380	
	acagcagcgg	ggacggcagc	caacgaatcc	tgtcggcctc	cgcgatctc	cacaggcagc	1440	
	ggcgctcccc	cgctcgacgt	gchgcttcg	cgccgcctcc	cttctcccg	acgcgtggc	1500	
25	gacgcgtgg	gchgacgcgt	ggg				1523	
	<210>	74						
	<211>	758						
	<212>	DNA						
30	<213>	Homo sapiens						
	<400>	74						
	gaattccggca	cgagacasgg	tttcacccctg	ttggccagga	tggtctcaat	ctcttgaccc	60	
	cgtgatctgc	ctgcctcggc	ctcccaaagt	gttaggatta	caggcatgag	ccactgtgcc	120	
35	cgccctttgt	tttttgagac	cttttttatt	ttgttgcac	ccaggctgaa	gtgcagtggc	180	
	acaaacacag	ttcactacag	ccttgcaccc	ctgggctcaa	gcaattctgc	ctcagtc	240	
	caagtaggtg	ggcttacaaa	tgcacagcat	gacacccggc	ttattttgc	attttgtgt	300	
	tgtgtgtgt	agccactg	caggccttgg	gcagcttct	tgtatctgt	tacctcatct	360	
40	ataaaaatgt	gataataata	gcttctccct	tattggggaa	ttgtatgtat	taaatgagat	420	
	aacatgtaaa	atgctcgt	caggccaggc	atggtggc	acgcttgc	tccca	480	
	ttgggaggct	gaggctgt	gatctttga	ggccagc	taagacc	ctggccaaata	540	
	ttgtgaaacc	ctgtgtctac	aaaaaaatac	agaaagt	ccaggcatgg	ttgtgc	600	
	ctgtggtccc	agctactc	aggctgaggt	gggagaatca	cttgg	ggagacagaa	660	
45	tttgaagtga	gccaagatgg	cgccactgca	ctctagcatg	ggctacag	tgagagc	720	
	tctcaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aactcgta			758	
	<210>	75						
	<211>	1096						
	<212>	DNA						
50	<213>	Homo sapiens						
	<400>	75						
	ccccacggct	cccatggc	tttcctgc	taccgtgtgg	aggcc	taac	ctgcgtggc	60
55	atcaatagct	tccgccc	caagtatgac	ctggtggc	tggcaagg	tttggagg	120	
	atgttccgca	agctcaacca	ccttc	ggag	cgcctgc	acc	agtcc	180
	ctcccccggc	tctcccg	ctgttccatc	ggcctctaca	tggccgt	ctgttcc	240	
	ctccctgg	ttgttctaa	ggctctgg	ctgtggat	gc	agctgat	ggctggaaat	300
	ggccttgagg	agccgggggg	tgccc	ttggc	ccc	actacagg	360	
60	gtggggctgg	cctcg	gtgt	ggcac	ctgt	ctgatctc	aggccatgg	420

	tatgtcctgc cagtgcgtggg ccaacacgtt gccaccaggc acttcccagt ggcagaggct	480
	gaggctgtgg tgcgtacact gctggcgatt tatgcagctg gcctggccct gccycacaat	540
	acccacccggg tgtaagcac acaggccccca gacaggggct ggatggact gaagctggta	600
5	gcctgtatct accttagact gcagctgggc tgcatcgccc tcaccaactt ctcaactggc	660
	ttcctgtgg ccaccacat ggtgcccact gtcgcgttg ccaagcctca tgggccccgg	720
	accctctatg ctgcctgtct ggtgctgacc agcccccggc ccacgctcct tggcagcctg	780
	ttcctgtggc gggagctgca ggaggcgcca ctgtcaactgg ccgagggctg gcagctttc	840
	ctggcagcgc tagcccaggg tgcgtggag caccacacta cgccgcctg ctcttccac	900
10	tgctgtccct gggcctctac cccgtctggc tgctttctg gaatgtgctc ttctggaaagt	960
	gagatctgcc tgcgtggct gggacagaga ctccccaaagg accccattct gcctccttct	1020
	ggggaaataaa atgagtgtct gtttcagcar maaaaaaaaaaaaaaa aaaaaaaaaaaa	1080
	aaaaaaaaagg gcggcc	1096
15	<210> 76	
	<211> 1230	
	<212> DNA	
	<213> Homo sapiens	
20	<400> 76	
	cacgagtgcc gctaacccttc ttcatccttt ggtggcaaag tagaaagatt ccagaattaa	60
	ctcgaccttt ctaaagacct gggctcagag gcagctggca ctgactgagc acccactatg	120
	tgccaggcac tgcgtgaat gcatttagatc atcaattatg aatttgacac caaggacctg	180
	gtgtgcctgg gcctgagcag catcggttgc gtctggtacc tgctgaggaa gcactggatt	240
25	gccaacaacc tttttggcct ggccttctcc cttaatggag tagagctcct gcacctaacc	300
	aatgtcagca ctggctgcatt cctgctggc ggactcttca tctacgtatgt ctctgggta	360
	tttggcacca atgtgtatggt gacagtggcc aagtccctcg aggccaccaat aaaattggtg	420
	tttcccagg atctgctgga gaaaggccctc gaagcaaaca actttgcct gctggactt	480
	ggagatgtcg tcattccagg gatcttcatt gccttgcgtc tgccgtttga catcagcttgc	540
30	aagaagaata cccacacacta cttctacacc agcttgcagc cttacatctt cggcctggc	600
	cttaccatct tcatcatgca catcttcaag catgctcagc ctgcctctt atacctggc	660
	cccgctgcgca tcggtttcc tgccctggtg gcgctggcca agggagaagt gacagagatg	720
	ttcagttatg aggagtcaaa tcctaaggat ccagccggcag tgacagaatc caaagaggg	780
	acagaggcat cagcatcgaa gggctggag aagaaagaga aatgtatgcag ctggtgcccg	840
35	agcctctcg gcccagacca gacagatggg ggctggccccc acacaggcgt gcacccgttag	900
	agggcacagg aggccaaggg cagctccagg acagggcagg gggcagcagg atacctccag	960
	ccaggccctct gtggcctctg ttcccttctc cctttcttgg ccctcctctg ctccctcccc	1020
	caccctgcag gaaaaagaaaa cccccagctt cccccctccc cgggagccag gtggaaaaag	1080
40	tgggtgtat ttttagattt tgcgtgtgg actgattttgc cttcacatta aaaactcatc	1140
	ccatggccag ggccggccac tgcgtctctg gaaaaaaaaaaaaaaa aaaaaaaaaaaa	1200
	aaaaaaaaaaa aaaaaaaaaaaa ggggaggggc	1230
45	<210> 77	
	<211> 911	
	<212> DNA	
	<213> Homo sapiens	
50	<400> 77	
	tcgaccacg cgtccgtctt cctaaaaggatg ccccttcca aagaaatttt aaaagaatct	60
	tatcaagggg ccctggagaa gaaaggatgt tgagggtcaa gtcacaactt tgagggaaaa	120
	tagaaagagg gctccttctt gagaaggaaatg aatttcaaaatg agtccaagag aaccaaaaat	180
	tcaggaccca ggagggtaaatg cattcctgtt tttgcgttcc tcaacagacca tttgagtgt	240
	tgggttttc aggtgacatt taaatgaaca aataatatcc atgtctcagg gtcagaatgt	300
55	gtactttgca actgattctg tccctcttgc gaggcttctg caagactgag aggggtggat	360
	gacttaatgtt acattaaaaaa caatgttatt aggckggata tggtggcaca tggcctgtat	420
	tctagcatt tgggargctg aggtggccag gcccargart tcaagaccatg tctggcaac	480
	atgggtgagac cctgttatcta ataaaaatatac aaaaatttttgc ccaggcatgg tggcacacac	540
60	ctggagtcccc agctactcag gagactgagg tggaaagatc acctgagctc aggaagtcga	600
	ggctgcgttg agcacaaggatt gcactactgc actctagcct acatggatag gagttagacc	660

	tgtttgaaaa acaaaaaaaca atcaaaaaca aaaaaaaaca acccacacaa tgttatTTT	720
	aaaatactga ggggagagaa gttggggaaa aaaaggaaaa acctaaaact ctccataatc	780
	ctaccatcaG aaaattacac taatgtgata agtgacttc tcccctctga atctccaatt	840
	ccattacttG tagtaaatat gaatcttatt ccacaaactc agacatgcaa aaaaaaaaaa	900
5	aaaggcgC c	911
	<210> 78	
	<211> 488	
10	<212> DNA	
	<213> Homo sapiens	
	<220>	
	<221> SITE	
15	<222> (324)	
	<223> n equals a,t,g, or c	
	<220>	
	<221> SITE	
20	<222> (438)	
	<223> n equals a,t,g, or c	
	<220>	
	<221> SITE	
25	<222> (484)	
	<223> n equals a,t,g, or c	
	<400> 78	
30	accgcagggg ctccccggacc ctgactctgc agccgaaccg gcacggTTTC gtggggacCC	60
	aggcttgcAA agtgacggTC attttcttT tctttctccc tcttgagtcc ttctgagatG	120
	atggctctGG ggcacggGA gctacccGGG tctttgtcGC gatggtagCG gcccGCTCTG	180
	gcccGCCACCC tctgtcGGGA gtgagcGCCA ccttgaACTC gttctCAAT tccaaACGCTA	240
	tcaagaacct gccccCACCG ctggggcGGG ctgcggggCA cccaaGGCTC tgcaGTCAGC	300
35	gcccGCGCCGG gaatctGTa CCCNGGGCGG gaataAGTAC cagaccATTG acaactacca	360
	gcccgtacCCS ttgcgcAAAA gaacraAAAGA aatttgcCGC actgaaATAA atttacttGc	420
	gcctaattCC ccacCCCNCC cgaaaAGGGG aaacCCCCGG ggcgttttC caaatttCttt	480
	tttnttCC	488
40	<210> 79	
	<211> 753	
	<212> DNA	
	<213> Homo sapiens	
45	<220>	
	<221> SITE	
	<222> (745)	
	<223> n equals a,t,g, or c	
50	<220>	
	<221> SITE	
	<222> (752)	
	<223> n equals a,t,g, or c	
55	<400> 79	
	gaattcggca cgagcggcgg ggtccatCC accccggta gcaggccaa ggcaGcGGGG	60
	gcccacacCC ctcacacgca aaactggCttt ttctggTCa ctgggtctG aaacccaaatC	120
	cagagcagCC tgtggctgt aaagcatata ttcttaatGA ctgcagactG gtgggatcat	180
	aggagccttC tgaatgacca ggactgctt ttggagat gatgaaaatG tactctttta	240
60	gcgtgttaga aatcaTTGt ttatTTGt ttctttggCC aagctggTC tagtgTTtCt	300



<221> SITE  
 <222> (5)  
 <223> n equals a,t,g, or c

5 <220>  
 <221> SITE  
 <222> (7)  
 <223> n equals a,t,g, or c

10 <220>  
 <221> SITE  
 <222> (9)  
 <223> n equals a,t,g, or c

15 <220>  
 <221> SITE  
 <222> (10)  
 <223> n equals a,t,g, or c

20 <220>  
 <221> SITE  
 <222> (1205)  
 <223> n equals a,t,g, or c

25 <400> 81
 

aaccnangnn	taccggtccg	gaattcccg	gtcggaccca	cgcgtccg	gcggggcgacg	60
cacgtcgagc	ggggggagccg	cgctgcctgt	ggagatccgc	ggaggccgac	aggattcg	120
ggctgcgc	ccccgtcg	tgcattgggt	taaaaaacgac	aaccaacatc	agccatgaaa	180
gatccaagt	gcagcgtac	tagcccaagc	atcatcaatg	aatgtgtat	tattaacgg	240
30 cattctcatg	aaatgacaa	tccatttgca	gagttatgt	ggatggaaaa	tgaagaagaa	300
ttcaacagac	aaatagaaga	ggagtttatgg	gaagaagaat	ttatttgcac	ctgtttccaa	360
gaaatgtcg	aaagggaaaga	agagcatgaa	tggtttatttc	cagctcgaga	tctcccacaa	420
actatggacc	aaatccaaga	ccagtttaat	gaccttgc	tcagtgtatgg	ctttctctg	480
35 gaagatctt	tggcaagag	caatctgaat	ccaaatgc	aggatgttgc	tcctgggg	540
aagtacggaa	atattttagt	agacggggcc	ctcttttgt	ggatgttagc	caatttccac	600
actgtgaagg	cagtatttgc	agacttattt	gtaaaagctc	tcttgcact	gttttacact	660
tatgcattgc	caaaatgttt	gttgcatttgc	catgcatttgc	aaaatgtctg	agactgttac	720
taagaaaaaa	gctgtcaaac	atttactgaa	aatagaatttgc	ccccatggc	ttgtatgtgaa	780
40 gacagcaagg	aaagaacac	cagtcaagtt	gtgacaacgc	accaaattaa	aagaccta	840
ccttacaaa	ttgtctttt	ttgaggctaa	tctatcactt	gttaatgtct	aaactttaaa	900
atcagtagat	ttaatttgc	ttccaactgt	taagcatatt	tctcagactt	aaatttgatt	960
atgtccccat	caaaaagaat	ctccatatttgc	tgaaggctgt	ttagttatttgc	ttagataatttgc	1020
45 tggtaaaggc	aagtatgtca	tattactgag	gctacaagg	agtcagcaga	tgagtgc	1080
tccagcctt	tctggatgt	tattgttgc	aatatttgc	tctaatgttgc	catctgaggr	1140
agtatgtat	tgagratgt	aacttctaa	gggttactg	catcatrgct	atgcctgtat	1200
ggrgnctwa	ccatatgacc	mataccamcc	cwtaatccca	gctgraccaa	rgrtacckgt	1260
aaccattwg	gatttgggg	gkggccttc	ccyggcytt	kttwacccmt	ccacggagaa	1320
	tctggca					1327
50	<210> 82					
	<211> 758					
	<212> DNA					
	<213> Homo sapiens					
55	<400> 82					
	gaattcggca	cgagacacgg	tttcacccctg	ttggccagga	tggtctcaat	60
	cgtgatctgc	ctgcctcg	ctcccaaagt	gctaggatta	caggcatgag	120
	cggcccttgc	tttttgc	ctttttatttgc	ttgttgcac	ccaggctgaa	180
60	acaaacacag	ttcactacag	ccttgaccc	ctgggctcaa	gcaatttgc	240

caagtaggtg ggcttacaaa tgcacagcat gacacctggc ttatTTTGTGtg	300
tgtgtgtgtg agccactgcg caggccttgg gcagcttct tgatctctgt tacctcatct	360
ataaaatgtat gataataata gcttctccct tattggggaa ttgtaatgtat taaaatgagat	420
5 aacatgtaaa atgctcaga caggccaggc atggcgctc acgcttgcaa tcccagcact	480
ttgggaggct gaggtgtcta gatctttaa ggccaggcagt taagaccagc ctggccaata	540
tggtaaacc ctgtgtctac caaaaaataac agaaaagtca ccaggcatgg tggtgcatgc	600
ctgtggtccc agtactcag aggtgaggt gggagaatca cttgagcccg ggagacagaa	660
gttgaagtga gccaagatgg cgccactgca ctctagcatg ggctacagag tgagagcctc	720
tctcaaaaaa aaaaaaaaaa aaaaaaaaaa aactcgta	758

10

&lt;210&gt; 83

&lt;211&gt; 48

&lt;212&gt; PRT

15 &lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (48)

20 &lt;223&gt; Xaa equals stop translation

&lt;400&gt; 83

Met Gly Ser Cys Ala Ala Phe Leu Leu Ala Ala Leu Ser Leu Leu Gly					
1	5		10		15
	10				
	15				

25

Val Leu Gly Gly Tyr Pro Gly Arg Arg Ala Phe Ile Leu Pro Asn Arg			
20	25		30
	30		

30

Arg Ser Leu Arg Gln Trp Leu Glu Val Ser Leu Gly Pro Val Ser Xaa			
35	40		45
	45		

35

&lt;210&gt; 84

&lt;211&gt; 38

&lt;212&gt; PRT

40 &lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (38)

45 &lt;223&gt; Xaa equals stop translation

&lt;400&gt; 84

Met Asn Glu Ala Pro Pro Leu Ser Ser Ser Ser Ile Cys Phe Ile Leu					
1	5		10		15
	10				
	15				

50

Phe Tyr Phe Phe Pro Leu Leu Pro Pro Leu Ser Ser Thr Cys Phe Ser			
20	25		30
	30		

Lys Gly Asn Arg His Xaa

35

55

&lt;210&gt; 85

&lt;211&gt; 53

&lt;212&gt; PRT

60 &lt;213&gt; Homo sapiens

<220>  
<221> SITE  
<222> (53)  
5 <223> Xaa equals stop translation

<400> 85  
Met Cys Gln Asn Arg Glu Ser Val Leu Val Leu Leu Ile Glu Ser Asn  
1 5 10 15

10 Met Phe Ser Phe Tyr Leu Leu Phe Ser Phe Tyr Ile Val Phe Ser Phe  
20 25 30

Phe Ile Val Leu Arg Pro Leu Pro Arg Asn Glu Ser Ile Lys Lys Ile  
15 35 40 45

Gly Val Ile Phe Xaa  
50

20 <210> 86  
<211> 26  
<212> PRT  
<213> Homo sapiens

25 <220>  
<221> SITE  
<222> (26)  
<223> Xaa equals stop translation

30 <400> 86  
Met Thr Val Leu Ala Lys Arg Leu Val Leu Phe Leu Gly His Ile Phe  
1 5 10 15

35 Leu Leu Leu Cys Val Arg Ile Leu Asp Xaa  
20 25

40 <210> 87  
<211> 78  
<212> PRT  
<213> Homo sapiens

45 <220>  
<221> SITE  
<222> (43)  
<223> Xaa equals any one of the naturally occurring L-amino acids

50 <220>  
<221> SITE  
<222> (78)  
<223> Xaa equals stop translation

<400> 87  
55 Met Ala Ala Arg Ser Ala Leu Ala Leu Leu Leu Leu Pro Val Leu  
1 5 10 15

Leu Leu Pro Val Gln Ser Arg Ser Glu Pro Glu Thr Thr Ala Pro Thr  
20 25 30

Pro Thr Pro Ile Pro Gly Gly Asn Ser Ser Xaa Ser Arg Pro Leu Pro  
35 40 45

Ser Ile Glu Leu His Ala Cys Gly Pro Tyr Pro Lys Pro Gly Leu Leu  
5 50 55 60

Ile Leu Leu Ala Pro Leu Ala Leu Trp Pro Ile Leu Leu Xaa  
65 70 75

10 <210> 88  
<211> 38  
<212> PRT  
<213> Homo sapiens

15 <220>  
<221> SITE  
<222> (38)  
<223> Xaa equals stop translation

20 <400> 88  
Met Cys Tyr Ile Pro Gly Ser Thr Gly Gly Gln Cys Trp Pro Trp Cys  
1 5 10 15

25 Trp Cys Trp Leu Cys Arg Glu Ala Leu Glu Trp Leu Cys Gly Ala Val  
20 25 30

Ser Ala Gly Pro Ala Xaa  
35

30 <210> 89  
<211> 44  
<212> PRT  
<213> Homo sapiens

35 <220>  
<221> SITE  
<222> (40)  
<223> Xaa equals any one of the naturally occurring L-amino acids

40 <220>  
<221> SITE  
<222> (44)  
<223> Xaa equals stop translation

45 <400> 89  
Met Leu Leu Arg Ile Ile His Leu Val Ile Phe Phe Ile Asn Phe Ser  
1 5 10 15

50 Thr Ser Val Val Ile Val His Tyr Asn Val Leu Asn Tyr Arg Cys Leu  
20 25 30

55 Leu Lys Cys Arg Cys Arg Val Xaa Lys Tyr Ser Xaa  
35 40

60 <210> 90  
<211> 60  
<212> PRT

<213> Homo sapiens

<220>

<221> SITE

5 <222> (60)

<223> Xaa equals stop translation

<400> 90

10 Met Gln Asn Cys Leu Gly Ser Leu Ile Pro Gly Val Leu Phe Ser Leu  
1 5 10 15

Leu Leu Leu Pro Ser Met Phe Asn Ile Ile Leu Thr Gln Ser Lys Tyr  
20 25 30

15 Gly Glu Asn Ser Tyr Pro Ala Cys Phe Tyr Ser Ser Ser Asn Phe Pro  
35 40 45

Val Ser Ala Ile Thr Phe Leu Val Gly Val Val Xaa  
50 55 60

20

<210> 91

<211> 55

<212> PRT

25 <213> Homo sapiens

<220>

<221> SITE

<222> (55)

30 <223> Xaa equals stop translation

<400> 91

Met Val Val Ile Val Leu Thr Ser Asn Val Cys Ile Cys Gly Tyr Val  
1 5 10 15

35 Val His Ser Ala Leu Ile Pro Arg Arg Gln Gly Leu Phe Leu Phe Leu  
20 25 30

40 Phe Leu Val Met Phe Tyr Phe Ser Ile Ala Phe Asn Arg Ile Thr Lys  
35 40 45

Gly Thr Leu Ser Ser Gln Xaa

50 55

45

<210> 92

<211> 51

<212> PRT

50 <213> Homo sapiens

<220>

<221> SITE

<222> (51)

55 <223> Xaa equals stop translation

<400> 92

Met Val Ala Gln Leu Val Gly Cys Val Val Ser Cys Leu Phe Val Leu  
1 5 10 15

60 Leu Arg Phe Leu Ile Ser Thr Phe Gly Ile Met Ser Phe Asn Gly Phe

20	25	30
----	----	----

Val Ile Phe Val Thr Val Leu Ala Ala Tyr Asn Phe Ser Ala Gly Ala	35	40
---	----	----

5

Phe Thr Xaa		45
-------------	--	----

50

10 <210> 93	<211> 156	<212> PRT
-------------	-----------	-----------

<213> Homo sapiens		
--------------------	--	--

15 <220>	<221> SITE	<222> (156)
----------	------------	-------------

<223> Xaa equals stop translation		
-----------------------------------	--	--

20 <400> 93	Met Trp Pro Gln Glu Ala Trp Val Cys Ile Leu Val Leu Leu Gly Thr	1 5 10 15
-------------	---	-----------

25 Arg Val Gly Leu Cys Val Gly Asp Ser Leu Ala Pro Gln Ala Ser Leu	20 25 30	
--	----------	--

Ser Tyr Cys Tyr Ile Leu Lys Val Pro Leu Arg Pro Lys Pro Leu Trp	35 40 45	
---	----------	--

30 Gln Leu Ser Asn Glu Ser Ile Cys Ser Glu Tyr Arg Val Glu Gly Gly	50 55 60	
--	----------	--

Gln Gly His Gln Glu Leu Arg Met Phe Leu Arg Leu Met Arg Pro Arg	65 70 75 80	
---	-------------	--

35 Tyr Trp Val His Gly Gly Pro Arg Ser Leu Cys Asp Ser Cys Ser Leu	85 90 95	
--	----------	--

40 Leu Pro Pro Cys Leu Asp Pro Ala Ser Ala Gln Lys Ala Asn Ser Leu	100 105 110	
--	-------------	--

Asp Ser Lys Gly Leu Pro Arg Pro Ile Ser Met Ser Cys Ser Cys Gln	115 120 125	
---	-------------	--

45 Leu Pro Val Pro Ser Leu Asp Leu Ser Ser Cys Leu Ala Pro Ser Leu	130 135 140	
--	-------------	--

Pro Thr Pro His Ile Phe Thr Asn Lys Arg Lys Xaa	145 150 155	
---	-------------	--

50 55 <210> 94	<211> 61	<212> PRT
----------------	----------	-----------

<213> Homo sapiens		
--------------------	--	--

60 <220>	<221> SITE	<222> (61)
----------	------------	------------

<223> Xaa equals stop translation		
-----------------------------------	--	--

<400> 94  
Met Ser His His Ala Arg Pro Tyr Lys Ala Phe Arg Ile Val Ser Cys  
1 5 10 15  
5 Tyr Phe Tyr Leu Phe Ile Ile Val Val Val Ile Ile Leu Leu Leu Tyr  
20 25 30  
10 Pro Ile Ser Gln Gly Trp His Val Ala Asn Ile Val Phe Leu Lys Asn  
35 40 45  
Ile Ser Asp His Ile Leu Val Leu Leu Lys Thr Phe Xaa  
50 55 60  
15 <210> 95  
<211> 71  
<212> PRT  
<213> Homo sapiens  
20 <220>  
<221> SITE  
<222> (71)  
<223> Xaa equals stop translation  
25 <400> 95  
Met Trp Phe Glu Ile Leu Pro Gly Leu Ser Val Met Gly Val Cys Leu  
1 5 10 15  
30 Leu Ile Pro Gly Leu Ala Thr Ala Tyr Ile His Arg Phe Thr Asn Gly  
20 25 30  
Gly Lys Glu Lys Arg Val Ala His Phe Gly Tyr His Trp Ser Leu Met  
35 40 45  
35 Glu Arg Asp Arg Arg Ile Ser Gly Val Asp Arg Tyr Tyr Val Ser Lys  
50 55 60  
40 Gly Leu Glu Asn Ile Asp Xaa  
65 70  
  
<210> 96  
<211> 37  
45 <212> PRT  
<213> Homo sapiens  
  
<220>  
<221> SITE  
50 <222> (37)  
<223> Xaa equals stop translation  
  
<400> 96  
Met Val Phe Leu Leu Leu Leu Phe Gly Phe Phe Phe Asp Gly Ser  
55 1 5 10 15  
Leu Arg Ser Pro Leu Leu Ile Ile His Leu Gly Pro Ala Pro Thr  
20 25 30  
60 Phe Leu Gln Ile Xaa

35

5       <210> 97  
      <211> 60  
      <212> PRT  
      <213> Homo sapiens

10      <220>  
      <221> SITE  
      <222> (60)  
      <223> Xaa equals stop translation

15      <400> 97  
      Met Leu Cys Gln Thr Ile Pro Leu Cys Asn Arg Leu His Ile Val Phe  
                1                   5                           10                   15  
      Met Ile Leu Ile Lys Leu Tyr Val Glu Thr Glu Cys Glu Val Lys Ser  
                20                   25                                   30  
20      Glu His Lys Lys Ile Met His Asp Glu Ile Ala Tyr His Phe Ile Gly  
                35                   40                           45  
      Tyr Leu Leu Cys Ile Tyr Thr Leu Arg Pro Leu Xaa  
25      50                   55                           60

30      <210> 98  
      <211> 44  
      <212> PRT  
      <213> Homo sapiens

35      <220>  
      <221> SITE  
      <222> (44)  
      <223> Xaa equals stop translation

40      <400> 98  
      Met Ser Val Ser Ser Asn Leu Trp Gln Thr Leu Ile Leu Leu Ser  
                1                   5                           10                   15  
      Leu Trp Phe Cys Leu Phe Pro Glu Cys His Ile Val Gly Ile Ile Gln  
                20                   25                           30  
45      Leu Cys Arg Leu Phe Arg Leu Pro Ser Phe Thr Xaa  
                35                   40

50      <210> 99  
      <211> 32  
      <212> PRT  
      <213> Homo sapiens

55      <220>  
      <221> SITE  
      <222> (32)  
      <223> Xaa equals stop translation

60      <400> 99  
      Met Cys Cys Arg Ala Gly Gly Ser Gln Ser Pro Gln Val Met Val Val

1	5	10	15
Leu Ile Ile Ile Leu Gly Pro Trp Gly Gly Val Arg Ile Asp Ala Xaa			
20	25	30	

5

10	<210> 100 <211> 180 <212> PRT <213> Homo sapiens		
15	<400> 100 Met Tyr Ser Cys Leu Leu Leu Pro Asp Leu Leu Tyr Leu Thr Leu Ser	1	5
		10	15
20	Pro Leu Val Val Ala Met Leu Leu Thr Pro His Phe Asn Val Ala Asn	20	25
			30
	Pro Gln Asn Leu Leu Ala Gly Leu Trp Leu Glu Asn Glu His Ser Phe	35	40
			45
25	Thr Leu Met Ala Pro Glu Arg Ala Arg Thr His His Cys Gln Pro Glu	50	55
			60
30	Glu Arg Lys Val Leu Phe Cys Leu Phe Pro Ile Val Pro Asn Ser Gln	65	70
			75
	Ala Gln Val Gln Pro Pro Gln Met Pro Pro Phe Cys Cys Ala Ala Ala	85	90
			95
35	Lys Glu Lys Thr Gin Glu Glu Gln Leu Gln Glu Pro Leu Gly Ser Gln	100	105
			110
	Cys Pro Asp Thr Cys Pro Asn Ser Leu Cys Pro Ser His Thr Gln Leu	115	120
			125
40	Thr Lys Ala Asn Thr Leu Ser Leu Phe Phe Phe Ser Phe Phe Leu	130	135
			140
	Ser Arg Val Ser Leu Leu Ser Pro Arg Leu Glu Cys Asn Gly Arg Ile	145	150
			155
45	Leu Ala His Cys Asn Leu His Leu Pro Gly Ser Ser Asn Ser Pro Val	165	170
			175
50	Ser Ala Ser Arg	180	
	<210> 101 <211> 212 <212> PRT <213> Homo sapiens		
55	<220>		
	<221> SITE		
60	<222> (45)		

<223> Xaa equals any one of the naturally occurring L-amino acids

<220>

<221> SITE

5 <222> (195)

<223> Xaa equals any one of the naturally occurring L-amino acids

<220>

<221> SITE

10 <222> (212)

<223> Xaa equals stop translation

<400> 101

Met Arg Leu Phe Leu Trp Asn Ala Val Leu Thr Leu Phe Val Thr Ser

15 1 5 10

15

Leu Ile Gly Ala Leu Ile Pro Glu Pro Glu Val Lys Ile Glu Val Leu  
20 25 30

20 Gln Lys Pro Phe Ile Cys His Arg Lys Thr Lys Gly Xaa Asp Leu Met  
35 40 45

Leu Val His Tyr Glu Gly Tyr Leu Glu Lys Asp Gly Ser Leu Phe His  
50 55 60

25 Ser Thr His Lys His Asn Asn Gly Gln Pro Ile Trp Phe Thr Leu Gly  
65 70 75 80

30 Ile Leu Glu Ala Leu Lys Gly Trp Asp Gln Gly Leu Lys Gly Met Cys  
85 90 95

Val Gly Glu Lys Arg Lys Leu Ile Ile Pro Pro Ala Leu Gly Tyr Gly  
100 105 110

35 Lys Glu Gly Lys Gly Lys Ile Pro Pro Glu Ser Thr Leu Ile Phe Asn  
115 120 125

Ile Asp Leu Leu Glu Ile Arg Asn Gly Pro Arg Ser His Glu Ser Phe  
130 135 140

40 Gln Glu Met Asp Leu Asn Asp Asp Trp Lys Leu Ser Lys Asp Glu Val  
145 150 155 160

Lys Ala Tyr Leu Lys Lys Glu Phe Glu Lys His Gly Ala Val Val Asn  
45 165 170 175

Glu Ser His His Asp Ala Leu Val Glu Asp Ile Phe Asp Lys Glu Asp  
180 185 190

50 Glu Asp Xaa Tyr Gly Phe Ile Ser Ala Arg Glu Phe Thr Tyr Lys His  
195 200 205

Asp Glu Leu Xaa  
210

55

<210> 102

<211> 621

<212> PRT

60 <213> Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (137)

5 &lt;223&gt; Xaa equals any one of the naturally occurring L-amino acids

&lt;400&gt; 102

Met Gly Leu Leu Ser Asp Pro Val Arg Arg Arg Ala Leu Ala Arg Leu  
1 5 10 1510 Val Leu Arg Leu Asn Ala Pro Leu Cys Val Leu Ser Tyr Val Ala Gly  
20 25 3015 Ile Ala Trp Phe Leu Ala Leu Val Phe Pro Pro Leu Thr Gln Arg Thr  
35 40 45Tyr Met Ser Glu Asn Ala Met Gly Ser Thr Met Val Glu Glu Gln Phe  
50 55 6020 Ala Gly Gly Asp Arg Ala Arg Ala Phe Ala Arg Asp Phe Ala Ala His  
65 70 75 80Arg Lys Lys Ser Gly Ala Leu Pro Val Ala Trp Leu Glu Arg Thr Met  
85 90 9525 Arg Ser Val Gly Leu Glu Val Tyr Thr Gln Ser Phe Ser Arg Lys Leu  
100 105 11030 Pro Phe Pro Asp Glu Thr His Glu Arg Tyr Met Val Ser Gly Thr Asn  
115 120 125Val Tyr Gly Ile Leu Arg Ala Pro Xaa Ala Ala Ser Thr Glu Ser Leu  
130 135 14035 Val Leu Thr Val Pro Cys Gly Ser Asp Ser Thr Asn Ser Gln Ala Val  
145 150 155 160Gly Leu Leu Leu Ala Leu Ala Ala His Phe Arg Gly Gln Ile Tyr Trp  
165 170 17540 Ala Lys Asp Ile Val Phe Leu Val Thr Glu His Asp Leu Leu Gly Thr  
180 185 19045 Glu Ala Trp Leu Glu Ala Tyr His Asp Val Asn Val Thr Gly Met Gln  
195 200 205Ser Ser Pro Leu Gln Gly Arg Ala Ile Gln Ala Ala Val Ala  
210 215 22050 Leu Glu Leu Ser Ser Asp Val Val Thr Ser Leu Asp Val Ala Val Glu  
225 230 235 240Gly Leu Asn Gly Gln Leu Pro Asn Leu Asp Leu Leu Asn Leu Phe Gln  
245 250 25555 Thr Phe Cys Gln Lys Gly Gly Leu Leu Cys Thr Leu Gln Gly Lys Leu  
260 265 27060 Gln Pro Glu Asp Trp Thr Ser Leu Asp Gly Pro Leu Gln Gly Leu Gln  
275 280 285

Thr Leu Leu Leu Met Val Leu Arg Gln Ala Ser Gly Arg Pro His Gly  
 290 295 300

5 Ser His Gly Leu Phe Leu Arg Tyr Arg Val Glu Ala Leu Thr Leu Arg  
 305 310 315 320

Gly Ile Asn Ser Phe Arg Gln Tyr Lys Tyr Asp Leu Val Ala Val Gly  
 325 330 335

10 Lys Ala Leu Glu Gly Met Phe Arg Lys Leu Asn His Leu Leu Glu Arg  
 340 345 350

15 Leu His Gln Ser Phe Phe Leu Tyr Leu Leu Pro Gly Leu Ser Arg Phe  
 355 360 365

Val Ser Ile Gly Leu Tyr Met Pro Ala Val Gly Phe Leu Leu Leu Val  
 370 375 380

20 Leu Gly Leu Lys Ala Leu Glu Leu Trp Met Gln Leu His Glu Ala Gly  
 385 390 395 400

Met Gly Leu Glu Glu Pro Gly Gly Ala Pro Gly Pro Ser Val Pro Leu  
 405 410 415

25 Pro Pro Ser Gln Gly Val Gly Leu Ala Ser Leu Val Ala Pro Leu Leu  
 420 425 430

Ile Ser Gln Ala Met Gly Leu Ala Leu Tyr Val Leu Pro Val Leu Gly  
 30 435 440 445

Gln His Val Ala Thr Gln His Phe Pro Val Ala Glu Ala Val  
 450 455 460

35 Val Leu Thr Leu Leu Ala Ile Tyr Ala Ala Gly Leu Ala Leu Pro His  
 465 470 475 480

Asn Thr His Arg Val Val Ser Thr Gln Ala Pro Asp Arg Gly Trp Met  
 485 490 495

40 Ala Leu Lys Leu Val Ala Leu Ile Tyr Leu Ala Leu Gln Leu Gly Cys  
 500 505 510

Ile Ala Leu Thr Asn Phe Ser Leu Gly Phe Leu Leu Ala Thr Thr Met  
 45 515 520 525

Val Pro Thr Ala Ala Leu Ala Lys Pro His Gly Pro Arg Thr Leu Tyr  
 530 535 540

50 Ala Ala Leu Leu Val Leu Thr Ser Pro Ala Ala Thr Leu Leu Gly Ser  
 545 550 555 560

Leu Phe Leu Trp Arg Glu Leu Gln Glu Ala Pro Leu Ser Leu Ala Glu  
 565 570 575

55 Gly Trp Gln Leu Phe Leu Ala Ala Leu Ala Gln Gly Val Leu Glu His  
 580 585 590

His Thr Tyr Gly Ala Leu Leu Phe Pro Leu Leu Ser Leu Gly Leu Tyr  
 60 595 600 605

Pro Cys Trp Leu Leu Phe Trp Asn Val Leu Phe Trp Lys  
 610 615 620

5 <210> 103  
 <211> 287  
 <212> PRT  
 <213> Homo sapiens

10 <220>  
 <221> SITE  
 <222> (263)  
 <223> Xaa equals any one of the naturally occurring L-amino acids

15 <400> 103  
 Met Ala Leu Leu Pro Ile Phe Phe Gly Ala Leu Arg Ser Val Arg Cys  
 1 5 10 15

20 Ala Arg Gly Lys Asn Ala Ser Asp Met Pro Glu Thr Ile Thr Ser Arg  
 20 25 30

Asp Ala Ala Arg Phe Pro Ile Ile Ala Ser Cys Thr Leu Leu Gly Leu  
 35 40 45

25 Tyr Leu Phe Phe Lys Ile Phe Ser Gln Glu Tyr Ile Asn Leu Leu Leu  
 50 55 60

30 Ser Met Tyr Phe Val Leu Gly Ile Leu Ala Leu Ser His Thr Ile  
 65 70 75 80

Ser Pro Phe Met Asn Lys Phe Phe Pro Ala Ser Phe Pro Asn Arg Gln  
 85 90 95

35 Tyr Gln Leu Leu Phe Thr Gln Gly Ser Gly Glu Asn Lys Glu Glu Ile  
 100 105 110

Ile Asn Tyr Glu Phe Asp Thr Lys Asp Leu Val Cys Leu Gly Leu Ser  
 115 120 125

40 Ser Ile Val Gly Val Trp Tyr Leu Leu Arg Lys His Trp Ile Ala Asn  
 130 135 140

Asn Leu Phe Gly Leu Ala Phe Ser Leu Asn Gly Val Glu Leu Leu His  
 45 145 150 155 160

Leu Asn Asn Val Ser Thr Gly Cys Ile Leu Leu Gly Gly Leu Phe Ile  
 165 170 175

50 Tyr Asp Val Phe Trp Val Phe Gly Thr Asn Val Met Val Thr Val Ala  
 180 185 190

Lys Ser Phe Glu Ala Pro Ile Lys Leu Val Phe Pro Gln Asp Leu Leu  
 195 200 205

55 Glu Lys Gly Leu Glu Ala Asn Asn Phe Ala Met Leu Gly Leu Gly Asp  
 210 215 220

Val Val Ile Pro Gly Ile Phe Ile Ala Leu Leu Leu Arg Phe Asp Ile  
 60 225 230 235 240

Ser Leu Lys Lys Asn Thr His Thr Tyr Phe Tyr Thr Ser Phe Ala Ala  
245 250 255

5 Tyr Ile Phe Gly Leu Gly Xaa Tyr His Leu His His Ala His Leu Gln  
260 265 270

Ala Cys Ser Val Met Arg Ser Gln Ile Leu Arg Ile Gln Arg Gln  
275 280 285

10

<210> 104  
<211> 32  
<212> PRT  
15 <213> Homo sapiens

<220>  
<221> SITE  
<222> (32)  
20 <223> Xaa equals stop translation

<400> 104  
Met Ser Arg Leu Leu Leu Phe Gly Arg Leu Cys Ser Leu Trp Cys  
1 5 10 15

25 Leu Ser Trp Leu Tyr Ser Thr Asp Thr Arg Pro Leu Leu Arg Gly Xaa  
20 25 30

30

<210> 105  
<211> 77  
35 <212> PRT  
<213> Homo sapiens

<400> 105  
Met Leu Pro Arg Leu Val Leu Asn Ser Trp Ala Cys Pro Pro Gln Pro  
40 1 5 10 15

Pro Lys Val Leu Glu Leu Gln Ala Cys Ala Thr Ile Ser Ser Leu Ile  
20 25 30

45 Thr Leu Phe Leu Met Phe Ile Lys Ser Ser His Pro Leu Ser Leu Ala  
35 40 45

Glu Ala Ser Gln Glu Gly Gln Asn Gln Leu Gln Ser Thr Ile Ser Asp  
50 55 60

50 Pro Glu Thr Trp Ile Leu Phe Val His Leu Asn Val Thr  
65 70 75

55 <210> 106  
<211> 45  
<212> PRT  
<213> Homo sapiens

60 <220>

<221> SITE  
<222> (45)  
<223> Xaa equals stop translation

5 <400> 106  
Met Val Phe Leu Val Phe Tyr Val Leu Arg Ala Leu Lys Cys Asn Ser  
1 5 10 15  
Ser Tyr His Ser Cys Thr Asn Val Leu Thr Gln Ile Ala Ser Gln Ile  
10 20 25 30  
Asp Lys Thr Leu Asn Asn Phe Ser Leu Lys Lys Cys Xaa  
35 40 45  
15 <210> 107  
<211> 42  
<212> PRT  
<213> Homo sapiens  
20 <220>  
<221> SITE  
<222> (42)  
<223> Xaa equals stop translation  
25 <400> 107  
Met Asn Pro Cys Leu Ser Ile Ile Phe Leu Leu Thr Pro Val Leu Leu  
1 5 10 15  
30 Ser His Pro Leu Gln Ser Leu His Phe Leu Leu Lys Val Asp Leu Asp  
20 25 30  
Phe Ser Leu Ser Cys Ser Ile Cys Thr Xaa  
35 35 40  
35 <210> 108  
<211> 70  
<212> PRT  
40 <213> Homo sapiens  
<220>  
<221> SITE  
<222> (70)  
45 <223> Xaa equals stop translation  
<400> 108  
Met Thr Val Tyr Leu Leu Lys Thr His Pro Cys Phe Phe Val Ala Tyr  
1 5 10 15  
50 Gln Met Gln Val Ala Leu Ile Ile Leu Leu Pro Gly Leu Arg Asn Ser  
20 25 30  
Lys Thr Val Thr Met Pro Leu Ser Pro Ala Leu Leu Pro Thr Leu Leu  
55 35 40 45  
Phe Phe Pro Ser Pro Thr Pro Phe Phe His Pro Phe Leu Ser Val Leu  
50 55 60  
60 Cys Cys Phe Lys Tyr Xaa

65                   70

5                   <210> 109  
       <211> 49  
       <212> PRT  
       <213> Homo sapiens

10                  <220>  
 10        <221> SITE  
       <222> (43)  
       <223> Xaa equals any one of the naturally occurring L-amino acids

15                  <220>  
 15        <221> SITE  
       <222> (49)  
       <223> Xaa equals stop translation

20                  <400> 109  
 20        Met His Ala Thr Cys Thr Arg Thr Trp Arg Ala Gln Val Ser Leu His  
                   1               5                   10                   15

25                  Gln Pro Pro Cys Ser Arg Asp Trp Lys Ile Cys His Leu Leu Val Val  
                   20               25                                   30

25                  Leu Ser Leu Pro Pro Pro Thr Pro Ala Arg Xaa Pro Glu Phe Leu Asn  
                   35               40                           45

30                  Xaa

35                  <210> 110  
       <211> 193  
       <212> PRT  
       <213> Homo sapiens

40                  <220>  
 40        <221> SITE  
       <222> (193)  
       <223> Xaa equals stop translation

45                  <400> 110  
 45        Met Ile Arg Asn Asp Gln Asp Ser Leu Met Gln Leu Leu Gln Leu Gly  
                   1               5                   10                   15

50                  Leu Val Val Leu Gly Ser Gln Glu Ser Gln Glu Ser Asp Leu Ser Lys  
                   20               25                           30

50                  Gln Leu Ile Ser Val Ile Ile Gly Leu Gly Val Ala Leu Leu Leu Val  
                   35               40                           45

55                  Leu Val Ile Met Thr Met Ala Phe Val Cys Val Arg Lys Ser Tyr Asn  
                   50               55                           60

55                  Arg Lys Leu Gln Ala Met Lys Ala Ala Lys Glu Ala Arg Lys Thr Ala  
                   65               70                           75                   80

60                  Ala Gly Val Met Pro Ser Ala Pro Ala Ile Pro Gly Thr Asn Met Tyr  
                   85               90                           95

Asn Thr Glu Arg Ala Asn Pro Met Leu Asn Leu Pro Asn Lys Asp Leu  
 100 105 110

5 Gly Leu Glu Tyr Leu Ser Pro Ser Asn Asp Leu Asp Ser Val Ser Val  
 115 120 125

Asn Ser Leu Asp Asp Asn Ser Val Asp Val Asp Lys Asn Ser Gln Glu  
 130 135 140

10 Ile Lys Glu His Arg Pro Pro His Thr Pro Pro Glu Pro Asp Pro Glu  
 145 150 155 160

Pro Leu Ser Val Val Leu Leu Gly Arg Gln Ala Gly Ala Ser Gly Gln  
 15 165 170 175

Leu Glu Gly Pro Ser Tyr Thr Asn Ala Gly Leu Asp Thr Thr Asp Leu  
 180 185 190

20 Xaa

25 <210> 111  
 <211> 71  
 <212> PRT  
 <213> Homo sapiens

30 <220>  
 <221> SITE  
 <222> (64)  
 <223> Xaa equals any one of the naturally occurring L-amino acids

35 <400> 111  
 Met Ala His Val Val Val Ala Arg Asn Glu Cys Leu Ile Arg Ala Phe  
 1 5 10 15

Leu Phe Leu Leu His Cys Val Ser Leu Leu Pro Ser Pro Gly Glu Val  
 20 25 30

40 Asn Ile Arg His Thr Leu Phe Thr Val Glu Glu Arg Leu Thr Thr Pro  
 35 40 45

45 Arg Ala Leu Lys Leu Ser Leu Ser Leu Ile Val Ser Leu His Ala Xaa  
 50 55 60

Cys Arg Lys Gln Glu Cys Ser  
 65 70

50 <210> 112  
 <211> 36  
 <212> PRT  
 <213> Homo sapiens

55 <220>  
 <221> SITE  
 <222> (36)  
 <223> Xaa equals stop translation

60

<400> 112  
Met Arg Leu Thr Glu Lys Asp Thr Val Leu Phe Thr Lys Gly Val Leu  
1 5 10 15

5 Phe Leu His Leu Phe Ile Asn Ala Leu Phe Trp Tyr Cys Lys Phe Gly  
20 25 30

His Asn Phe Xaa  
35

10

<210> 113  
<211> 60  
<212> PRT  
15 <213> Homo sapiens

<220>  
<221> SITE  
<222> (60)  
20 <223> Xaa equals stop translation

<400> 113  
Met Thr Ser Val Ser Thr Gln Leu Ser Leu Val Leu Met Ser Leu Leu  
1 5 10 15

25 Leu Val Leu Pro Val Val Glu Ala Val Glu Ala Gly Asp Ala Ile Ala  
20 25 30

Leu Leu Leu Gly Val Val Leu Ser Ile Thr Gly Ile Cys Ala Cys Leu  
30 35 40 45

Gly Val Tyr Ala Arg Lys Arg Asn Gly Gln Met Xaa  
50 55 60

35

<210> 114  
<211> 29  
<212> PRT  
40 <213> Homo sapiens

<220>  
<221> SITE  
<222> (29)  
45 <223> Xaa equals stop translation

<400> 114  
Met Asn Ser Phe Trp Ser Lys Leu Leu Val Leu Pro Leu Leu Ala Pro  
1 5 10 15

50 Leu Ser Met Ala Arg Ala Ser Ala Cys Gln Arg Trp Xaa  
20 25

55 <210> 115  
<211> 25  
<212> PRT  
<213> Homo sapiens

60 <220>  
<221> SITE

&lt;222&gt; (25)

&lt;223&gt; Xaa equals stop translation

&lt;400&gt; 115

5 Met Met Arg Leu Leu Asp Leu Arg Ile Phe Leu Met Ile His His Lys  
1 5 10 15

Ala Lys Ser Trp Glu Ser His Thr Xaa  
20 25

10

&lt;210&gt; 116

&lt;211&gt; 35

&lt;212&gt; PRT

15 <213> Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (35)

20 <223> Xaa equals stop translation

&lt;400&gt; 116

Met Pro Leu Ser Leu Leu Leu Ile Val Trp Lys Leu Glu Leu Cys Val  
1 5 10 15

25

Gly Ser Ala Leu Val Leu Ile His Thr Gln Arg Arg Tyr Ile Ile Leu  
20 25 30

30 Gln Val Xaa  
35

&lt;210&gt; 117

&lt;211&gt; 78

35 <212> PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (78)

40 <223> Xaa equals stop translation

&lt;400&gt; 117

Met Leu Leu Ala Thr Leu Leu Leu Leu Leu Gly Gly Ala Leu Ala  
45 1 5 10 15

His Pro Asp Arg Ile Ile Phe Pro Asn His Ala Cys Glu Asp Pro Pro  
20 25 30

50 Ala Val Leu Leu Glu Val Gln Gly Thr Leu Gln Arg Pro Leu Val Arg  
35 40 45

Asp Ser Arg Thr Ser Pro Ala Asn Cys Thr Trp Leu Thr Lys Arg Val  
50 55 60

55 Gln Gln Met Leu Leu Phe His Ser Tyr Gly Ile Ala Gln Xaa  
65 70 75

60 <210> 118

<211> 44  
<212> PRT  
<213> Homo sapiens

5      <220>  
      <221> SITE  
      <222> (44)  
      <223> Xaa equals stop translation

10     <400> 118  
Met Thr Gly Val Phe Lys Leu Pro Leu Leu Phe Trp Val His Glu Ala  
      1                  5                  10                  15  
Ser Val Gly Gly Cys Pro Tyr Val Lys Leu Val Glu Phe Glu Glu Met  
15     20                  25                  30  
Leu Thr Leu Tyr Gly Ile Leu Leu Ile Leu Phe Xaa  
      35                  40

20     <210> 119  
<211> 46  
<212> PRT  
<213> Homo sapiens

25     <220>  
      <221> SITE  
      <222> (46)  
      <223> Xaa equals stop translation

30     <400> 119  
Met Gln Leu Ala Pro Phe Ile Ser Ile Pro Val Leu Ser Gly Thr Thr  
      1                  5                  10                  15

35     Pro Trp Thr Ala Val Phe Arg Ala Ser Ser Ile Cys Thr Pro Leu Leu  
      20                  25                  30  
Thr Leu Ser Ala Ala Gly Met Leu Val Glu Ser Ser Leu Xaa  
      35                  40                  45

40     <210> 120  
<211> 29  
<212> PRT  
45     <213> Homo sapiens

50     <220>  
      <221> SITE  
      <222> (29)  
      <223> Xaa equals stop translation

55     <400> 120  
Met Pro Pro Leu Ser Asp Ile Leu Leu Thr Val Ala Val Val Ala Phe  
      1                  5                  10                  15  
Glu Met Thr Gly His Ile Tyr Ile Trp Pro His Thr Xaa  
      20                  25

60     <210> 121

<211> 62  
<212> PRT  
<213> Homo sapiens

5 <400> 121  
Met Glu Leu Pro Cys Asp Cys Ser Lys Leu Leu Tyr Cys Lys Phe Ser  
1 5 10 15

10 Val Trp His Leu Pro Val Asn Ala Met Lys Leu Leu Ile Ile Phe Leu  
20 25 30

Lys Val Leu His Cys Leu Phe Phe Leu Leu Cys Lys Phe Leu Tyr  
35 40 45

15 Thr Leu Ile Val Ile Leu Thr Asp Lys Tyr Ser Ile Leu Asn  
50 55 60

20 <210> 122  
<211> 87  
<212> PRT  
<213> Homo sapiens

25 <220>  
<221> SITE  
<222> (68)  
<223> Xaa equals any one of the naturally occurring L-amino acids

30 <220>  
<221> SITE  
<222> (72)  
<223> Xaa equals any one of the naturally occurring L-amino acids

35 <220>  
<221> SITE  
<222> (87)  
<223> Xaa equals stop translation

40 <400> 122  
Met Pro Val Ser Trp Gly Cys Pro Ser Lys Thr Pro Gln Thr Arg Ala  
1 5 10 15

Tyr Thr Arg Cys Val Tyr Phe Leu Met Val Leu Glu Ala Gly Val Gly  
20 25 30

45 Gly His Ser Val Ser Arg Val Gly Ser Leu Glu Val Pro Pro Trp Leu  
35 40 45

50 Val Ala Ala Asn Asn Phe Pro His Leu Met Trp Ser Ser Phe Cys Val  
50 55 60

Gly Pro His Xaa Val Phe Leu Xaa Asp Pro Ser Leu Pro Asp Pro Gly  
65 70 75 80

55 Pro Pro Asn Asn Leu Thr Xaa  
85

60 <210> 123  
<211> 64

<212> PRT  
<213> Homo sapiens

&lt;220&gt;

5 <221> SITE  
<222> (64)  
<223> Xaa equals stop translation

&lt;400&gt; 123

10 Met Cys Tyr Phe Leu Glu Ile Ser Leu Leu Met Val Phe Ala Leu Asn  
1 5 10 15Ile Lys Ala Ala Tyr Gly Cys Cys Asn Ile Asn Gly Thr Glu Val His  
20 25 3015 Arg Ala Lys Gly Pro Val Ser Val Pro Phe Pro Leu Ser Arg Pro Leu  
35 40 4520 Ser Gly Thr Pro Leu Leu Asp Arg Leu Arg Pro Phe Gln Thr Leu Xaa  
50 55 60

25

<210> 124  
<211> 36  
<212> PRT  
<213> Homo sapiens

30 <220>  
<221> SITE  
<222> (36)  
<223> Xaa equals stop translation35 <400> 124  
Met Pro Leu Pro Ser Ser Phe Pro Leu Pro Val Phe Leu Ser Ser Cys  
1 5 10 1540 Pro Phe Leu Met Ser Val Ser Ile Gly Phe Leu Ile Leu Val Phe Asn  
20 25 30Val His Pro Xaa  
3545 <210> 125  
<211> 32  
<212> PRT  
<213> Homo sapiens50 <220>  
<221> SITE  
<222> (32)  
<223> Xaa equals stop translation

&lt;400&gt; 125

55 Met Phe Ile Phe Cys Val Ser Leu Ala Phe Leu Pro Arg Phe Ile Ser  
1 5 10 15

60

Pro Gln Ser Cys Glu Trp Ala Gly Leu Ser Leu Val Trp His His Xaa  
20 25 30

5

<210> 126

<211> 41

10 <212> PRT

<213> Homo sapiens

<220>

<221> SITE

15 <222> (41)

<223> Xaa equals stop translation

<400> 126

Met Lys Asn Asn Thr Gln Lys Arg Leu Phe Leu Trp Gly Glu Leu Leu  
20 1 5 10 15

Leu Gln Asp Leu Ala Leu Ile Leu Tyr Leu Ser Ile Phe Leu Lys Ser  
20 25 30

25 Thr Leu Thr Asn Leu Asn Leu Phe Xaa  
35 40

<210> 127

30 <211> 28

<212> PRT

<213> Homo sapiens

<220>

35 <221> SITE

<222> (28)

<223> Xaa equals stop translation

<400> 127

40 Met Leu Asn Val Phe Phe Ser Leu Ile Leu Phe Phe Ser Pro Asn Arg  
1 5 10 15

Ala Leu Pro Ala Ile Ser Ser Cys Ile Thr Phe Xaa  
20 25

45

<210> 128

<211> 69

<212> PRT

50 <213> Homo sapiens

<220>

<221> SITE

<222> (69)

55 <223> Xaa equals stop translation

<400> 128

Met Arg Ala Val Gly Glu Arg Leu Leu Leu Lys Leu Gln Arg Leu Pro  
1 5 10 15

60

Gln Ala Glu Pro Val Glu Ile Val Ala Phe Ser Val Ile Ile Leu Phe  
           20                 25                 30

5    Thr Ala Thr Val Leu Leu Leu Leu Ile Ala Cys Ser Cys Cys Cys  
       35                 40                 45

    Thr His Cys Cys Cys Pro Glu Arg Arg Gly Arg Lys Val Gln Val Gln  
       50                 55                 60

10   Pro Thr Pro Pro Xaa  
       65

15   <210> 129  
   <211> 87  
   <212> PRT  
   <213> Homo sapiens

20   <400> 129  
   Met Asp Pro Arg Arg Val Thr Ala Cys Cys His Val Trp Thr Val Gly  
       1                 5                 10                 15

    Leu Phe Cys Ile Trp Ala Val Gly Leu Ser Cys Ser Leu Ser Leu Ser  
       20                 25                 30

25   His Val Ile Val Trp Leu Ser Gly Ala Gly Cys Thr Leu Ile Cys Glu  
       35                 40                 45

30   Asp Asn Pro Phe Leu Leu Leu Phe Ser Gln Tyr Leu Gln Pro His His  
       50                 55                 60

    Pro Glu Ile Met Lys Pro Phe Ile Leu Gly His Lys Ser Ser Asn Gly  
       65                 70                 75                 80

35   Gly Leu Ser Pro Pro Ser Ala  
       85

40   <210> 130  
   <211> 64  
   <212> PRT  
   <213> Homo sapiens

45   <220>  
   <221> SITE  
   <222> (64)  
   <223> Xaa equals stop translation

50   <400> 130  
   Met Phe Tyr Met Val Cys Val Leu Gly Ser Gly Ala Gln Pro Leu Ser  
       1                 5                 10                 15

    Glu Leu Ala Tyr Leu Ala Lys Leu Pro Thr Leu Gln Val Gly Lys Tyr  
       20                 25                 30

55   Asn Pro Leu Phe Asn Lys Ala His Pro Leu His Pro Val Leu Thr Thr  
       35                 40                 45

60   Phe Cys Glu Cys Ala Val Ile Phe Ser Cys Ser Ile Ala Arg Trp Xaa  
       50                 55                 60

5           <210> 131  
  <211> 54  
  <212> PRT  
  <213> Homo sapiens

10          <400> 131  
Met Arg Phe Gln Ser Tyr Leu Trp Pro Ser Arg Ile Leu Val Gly Thr  
  1               5               10               15

15          Tyr Cys Ile Ala Ala Glu Val Leu Phe Pro Ser Ala Leu Ala Ser Cys  
  20               25               30  
Gly Pro Val Trp Gln Gly Gly Ala Pro Thr Lys Ser Trp Gln Pro Gly  
  35               40               45

20          Ala Lys Thr Ile Ile Pro  
  50

25          <210> 132  
  <211> 41  
  <212> PRT  
  <213> Homo sapiens

30          <220>  
  <221> SITE  
  <222> (41)  
  <223> Xaa equals stop translation

35          <400> 132  
Met Arg Arg Trp Ala Gly Phe Gly Lys Ser Pro Gln Phe Trp Trp Thr  
  1               5               10               15

40          Gly Ile Leu Val Ala Leu Gly Ala Ala Leu Leu Gly Gly Pro Arg Leu  
  20               25               30

45          Gly Arg Arg Leu Thr Phe Gly Leu Xaa  
  35               40

50          <210> 133  
  <211> 69  
  <212> PRT  
  <213> Homo sapiens

55          <220>  
  <221> SITE  
  <222> (69)  
  <223> Xaa equals stop translation

60          <400> 133  
Met Ala Leu Ala Ile Phe Ile Pro Val Leu Ile Ile Ser Leu Leu Leu  
  1               5               10               15

65          Gly Gly Ala Tyr Ile Tyr Ile Thr Arg Cys Arg Tyr Tyr Ser Asn Leu

20 25 30

Arg Leu Pro Leu Met Tyr Ser His Pro Tyr Ser Gln Ile Thr Val Glu  
35 40 45

5 Thr Glu Phe Asp Asn Pro Ile Tyr Glu Thr Gly Glu Thr Arg Glu Tyr  
50 55 60

Glu Val Ser Ile Xaa  
10 65

<210> 134

<211> 48

15 <212> PRT

<213> Homo sapiens

<220>

<221> SITE

20 <222> (48)

<223> Xaa equals stop translation

<400> 134

Met Gly Phe Leu Phe Leu His Ile Leu Pro Ser Ile Ile Asn Thr Arg  
25 1 5 10 15

Ser Ala Pro Gln Pro Thr Ser Cys Arg Met Gln Pro Glu Gln Gln Pro  
20 25 30

30 His Ser Thr Leu Lys Pro Val Ile Leu Gly Met Met Ile Ile Ser Xaa  
35 40 45

35

<210> 135

<211> 76

40 <212> PRT

<213> Homo sapiens

<400> 135

Met Ser Gly Leu Val Gly Gly Ser Arg Cys Ser Lys Val Arg Phe  
1 5 10 15

45 Arg Cys Phe Asn Gly Asp Ser Leu Leu Val Leu Val Leu Gln His His  
20 25 30

50 Phe Arg Leu Cys Ser Trp Cys Leu Ala Pro Ser Leu Phe Leu Leu Leu  
35 40 45

Ser Cys Gln Val Val Ser Thr Met Met Glu Gln Asp Pro Val Ile Tyr  
50 55 60

55 Asp Asp Asp Asp Asp Asp Leu Pro Asn Tyr Phe Ser Val  
65 70 75

60 <210> 136

<211> 55

<212> PRT  
<213> Homo sapiens

5       <220>  
      <221> SITE  
      <222> (32)  
      <223> Xaa equals any one of the naturally occurring L-amino acids

10      <220>  
      <221> SITE  
      <222> (39)  
      <223> Xaa equals any one of the naturally occurring L-amino acids

15      <220>  
      <221> SITE  
      <222> (55)  
      <223> Xaa equals stop translation

20      <400> 136  
      Met Phe Leu Glu Leu Pro Met Gln His Ser Asp Val Leu Leu Phe Leu  
          1             5                 10                 15

25      Val Cys Trp Lys Ala Met Gly Ser Lys Lys Ser Pro Ser His Phe Xaa  
          20                 25                 30

      Pro Glu Val Gly Gly Ile Xaa Pro Ser Phe Gly Met Leu Asn Val Thr  
          35                 40                 45

30      Leu Leu Arg Ser Leu Thr Xaa  
          50                 55

35      <210> 137  
      <211> 54  
      <212> PRT  
      <213> Homo sapiens

40      <400> 137  
      Met Leu Val Leu Phe Pro Leu Leu Tyr Arg Gly Trp Ser Pro Val Pro  
          1              5                 10                 15

45      Gly Thr Ala Glu Gly Gly Met Cys Cys Cys Cys Leu Cys Ile Ser Arg  
          20                 25                 30

      Tyr Ser Leu Leu Thr Ser Ser Gln Asp Lys Glu Pro Pro Tyr Glu Met  
          35                 40                 45

50      Ser Ser Ser Glu Leu Ser  
          50

55      <210> 138  
      <211> 36  
      <212> PRT  
      <213> Homo sapiens

60      <220>  
      <221> SITE  
      <222> (33)  
      <223> Xaa equals any one of the naturally occurring L-amino acids

<220>  
 <221> SITE  
 <222> (36)  
 5 <223> Xaa equals stop translation  
  
 <400> 138  
 Met Thr Cys Tyr Glu Val Ile Leu Phe Phe Ile Lys Leu Phe Ser Asp  
 1 5 10 15  
 10 Met Gly Lys Tyr Lys Glu Cys Lys Glu Phe Lys Lys Gln Arg Thr Lys  
 20 25 30  
  
 15 Xaa Tyr Met Xaa  
 35  
  
 <210> 139.  
 <211> 80  
 20 <212> PRT  
 <213> Homo sapiens  
  
 <400> 139  
 Met Lys Ala Gln Pro Leu Glu Ala Leu Leu Leu Val Ala Leu Val Leu  
 25 1 5 10 15  
  
 Ser Phe Cys Gly Val Trp Phe Glu Asp Trp Leu Ser Lys Trp Arg Phe  
 20 25 30  
  
 30 Gln Cys Ile Phe Gln Leu Ala His Gln Pro Ala Leu Val Asn Ile Gln  
 35 35 40 45  
  
 Phe Arg Gly Thr Val Leu Gly Ser Glu Thr Phe Leu Gly Ala Glu Glu  
 50 50 55 60  
 35 Asn Ser Ala Asp Val Arg Ser Trp Gln Thr Leu Ser Tyr Phe Glu Leu  
 65 65 70 75 80  
  
 40  
  
 <210> 140  
 <211> 67  
 45 <212> PRT  
 <213> Homo sapiens  
  
 <400> 140  
 Met Ala Ala Ser Val Gly Arg Ala Thr Arg Ser Ala Ala Ala His Leu  
 50 1 5 10 15  
  
 Thr Gln Leu Pro Pro Ala Pro Arg Ala Gln Arg Thr Ser Pro Ala Gln  
 20 25 30  
  
 55 Pro Asp Glu Gly Lys Arg Arg Asp Ala Asp Pro Trp Arg Thr Gly Pro  
 35 40 45  
  
 Thr Val Asn Lys Thr Gly Ser Ile Pro Gly Arg Leu Arg Gly Trp Ala  
 50 55 60  
 60

Arg Ala Glu  
65

5      <210> 141  
      <211> 51  
      <212> PRT  
      <213> Homo sapiens

10     <220>  
      <221> SITE  
      <222> (51)  
      <223> Xaa equals stop translation

15     <400> 141  
Met Gly Trp Leu Cys Cys Glu Pro Ser Gly Leu Tyr Asn Leu Glu Lys  
      1              5                  10                  15

20     Gln Tyr Phe Phe Ser Ser Leu Gln Ala Gly Leu Pro Val Ile Val  
      20                                  25                  30

25     Ser Ser Gly Cys Thr Lys Ile Ala Tyr Gly Phe Ala Val Tyr Ser Pro  
      35                                  40                  45

30     Ser Ser Xaa  
      50

35     <210> 142  
      <211> 54  
      <212> PRT  
      <213> Homo sapiens

40     <400> 142  
Met Arg Arg Cys Val Arg His Val Leu Gly Ile Gly Leu Ile Val Leu  
      1              5                  10                  15

45     Lys Asn Leu Tyr Phe His Lys Asn Ser Met Tyr Pro Ser Pro Lys Leu  
      20                                  25                  30

50     Ser Ser Phe Gln Glu Ala Phe Leu Phe Phe Leu Ile Leu Lys Asn  
      35                                  40                  45

55     Pro Leu Thr Leu Cys Ser  
      50

60     <210> 143  
      <211> 50  
      <212> PRT  
      <213> Homo sapiens

65     <220>  
      <221> SITE  
      <222> (50)  
      <223> Xaa equals stop translation

70     <400> 143  
Ile His Pro Ser Arg Ser Thr Leu Ser Ser Gln Leu Val Thr Leu Pro  
      1              5                  10                  15

Leu Phe Glu Leu Val Phe Pro Ile Pro Ser Ser Gln Ser Pro Phe Ser  
20 25 30

5 Leu Asn Tyr Leu Ser Glu Phe Pro Leu Pro Glu His Glu Pro Cys Leu  
35 40 45

Glu Xaa  
50

10

<210> 144  
<211> 87  
<212> PRT  
15 <213> Homo sapiens

<220>  
<221> SITE  
<222> (84)  
20 <223> Xaa equals any one of the naturally occurring L-amino acids

<220>  
<221> SITE  
<222> (87)  
25 <223> Xaa equals stop translation

<400> 144  
Met Thr Cys Cys Cys Leu Leu Cys Lys Leu Gln Gly Ile Phe Phe  
1 5 10 15

30 Ser Phe Asn Ser Ser Val Leu Lys Ser Ile Leu Gly Thr Thr Arg Thr  
20 25 30

Leu Ser Ala Pro Trp Ile Gly Val Ser Val Lys Gly Thr Gln Trp Ala  
35 35 40 45

Leu Gly Ser Ala Arg Pro Gly Cys Gly Ser Gln Leu Thr Ser Ser Leu  
50 55 60

40 Gly Gly Leu Arg Gln Val Ile Cys Gln Pro His Leu Gln Lys His Asp  
65 70 75 80

Ala Lys Leu Xaa Ser Val Xaa  
85

45

<210> 145  
<211> 57  
<212> PRT  
50 <213> Homo sapiens

<400> 145  
Met His Lys Cys Asn Thr Val Thr Arg Glu Leu Leu Gln Leu Ser Leu  
1 5 10 15

55 Leu Ile Leu Pro Ser Gln Cys Gly Asn Cys Ala Thr Ser Thr Lys Arg  
20 25 30

Gly Pro Arg Leu Leu Lys Tyr Phe Arg Thr Ser Pro Gln Glu Gln Thr  
60 35 40 45

Pro Leu His Leu Asp Ser Asp Cys Ser  
50 55

5 <210> 146  
<211> 87  
<212> PRT  
<213> Homo sapiens

10 <400> 146  
Met Ser His Cys Ala Arg Pro Leu Phe Phe Glu Thr Phe Phe Ile Leu  
1 5 10 15

15 Leu Ser Pro Arg Leu Lys Cys Ser Gly Thr Asn Thr Val His Tyr Ser  
20 25 30

Leu Asp Leu Leu Gly Ser Ser Asn Ser Ala Ser Val Pro Gln Val Gly  
35 40 45

20 Gly Leu Thr Asn Ala Gln His Asp Thr Trp Leu Ile Phe Val Phe Cys  
50 55 60

25 Val Cys Val Cys Glu Pro Leu Arg Arg Pro Trp Ala Ala Phe Leu Ile  
65 70 75 80

Ser Val Thr Ser Ser Ile Lys  
85

30 <210> 147  
<211> 230  
<212> PRT  
<213> Homo sapiens

35 <220>  
<221> SITE  
<222> (216)  
<223> Xaa equals any one of the naturally occurring L-amino acids

40 <400> 147  
Met Gly Leu Ala Leu Tyr Val Leu Pro Val Leu Gly Gln His Val Ala  
1 5 10 15

45 Thr Gln His Phe Pro Val Ala Glu Ala Glu Ala Val Val Leu Thr Leu  
20 25 30

Leu Ala Ile Tyr Ala Ala Gly Leu Ala Leu Pro His Asn Thr His Arg  
35 40 45

50 Val Val Ser Thr Gln Ala Pro Asp Arg Gly Trp Met Ala Leu Lys Leu  
50 55 60

55 Val Ala Leu Ile Tyr Leu Ala Leu Gln Leu Gly Cys Ile Ala Leu Thr  
65 70 75 80

Asn Phe Ser Leu Gly Phe Leu Leu Ala Thr Thr Met Val Pro Thr Ala  
85 90 95

60 Ala Leu Ala Lys Pro His Gly Pro Arg Thr Leu Tyr Ala Ala Leu Leu

	100	105	110
	Val Leu Thr Ser Pro Ala Ala Thr Leu Leu Gly Ser Leu Phe Leu Trp		
	115	120	125
5	Arg Glu Leu Gln Glu Ala Pro Leu Ser Leu Ala Glu Gly Trp Gln Leu		
	130	135	140
10	Phe Leu Ala Ala Leu Ala Gln Gly Val Leu Glu His His Thr Thr Ala		
	145	150	155
	Pro Cys Ser Ser His Cys Cys Pro Trp Ala Ser Thr Pro Ala Gly Cys		
	165	170	175
15	Phe Ser Gly Met Cys Ser Ser Gly Ser Glu Ile Cys Leu Ser Gly Leu		
	180	185	190
	Gly Gln Arg Leu Pro Lys Asp Pro Ile Leu Pro Pro Ser Gly Glu Ile		
	195	200	205
20	Asn Glu Cys Leu Phe Gln Gln Xaa Lys Lys Lys Lys Lys Lys Lys		
	210	215	220
	Lys Lys Lys Gly		
25	225	230	
	<210> 148		
	<211> 63		
30	<212> PRT		
	<213> Homo sapiens		
	<220>		
	<221> SITE		
35	<222> (63)		
	<223> Xaa equals stop translation		
	<400> 148		
	Gln Pro Ala Leu Leu Tyr Leu Val Pro Ala Cys Ile Gly Phe Pro Val		
40	1	5	10
	15		
	Leu Val Ala Leu Ala Lys Gly Glu Val Thr Glu Met Phe Ser Tyr Glu		
	20	25	30
45	Glu Ser Asn Pro Lys Asp Pro Ala Ala Val Thr Glu Ser Lys Glu Gly		
	35	40	45
	Thr Glu Ala Ser Ala Ser Lys Gly Leu Glu Lys Lys Glu Lys Xaa		
	50	55	60
50	<210> 149		
	<211> 18		
	<212> PRT		
55	<213> Homo sapiens		
	<220>		
	<221> SITE		
	<222> (18)		
60	<223> Xaa equals stop translation		

<400> 149  
Gln Leu Ile Leu Ser Leu Leu Arg Gly Phe Cys Lys Thr Glu Arg Val  
1 5 10 15

5 Gly Xaa

10 <210> 150  
<211> 16  
<212> PRT  
<213> Homo sapiens

15 <220>  
<221> SITE  
<222> (16)  
<223> Xaa equals stop translation

20 <400> 150  
Met Ala Leu Gly Ala Arg Glu Leu Pro Gly Ser Leu Ser Arg Trp Xaa  
1 5 10 15

25

30 <210> 151  
<211> 22  
<212> PRT  
<213> Homo sapiens

35 <220>  
<221> SITE  
<222> (22)  
<223> Xaa equals stop translation

40 <400> 151  
Met Tyr Ser Phe Ser Val Leu Glu Ile Thr Cys Phe Ile Leu Phe Leu  
1 5 10 15

Trp Pro Ser Trp Val Xaa  
20

45 <210> 152  
<211> 25  
<212> PRT  
<213> Homo sapiens

50 <220>  
<221> SITE  
<222> (25)  
<223> Xaa equals stop translation

55 <400> 152  
Met Lys Ile Lys Gln Arg Phe Ser Leu Leu Leu Phe His Cys Pro Phe  
1 5 10 15

60 Pro Pro Cys Cys Leu Ser Leu Gly Xaa

20

25

5           <210> 153  
      <211> 40  
      <212> PRT  
      <213> Homo sapiens

10          <400> 153  
      Met Asn Gly Leu Phe Gln Leu Glu Ile Ser His Lys Leu Trp Thr Lys  
          1               5                   10               15  
          Ser Lys Thr Ser Leu Met Thr Leu Leu Ser Val Met Ala Leu Leu Trp  
          20              25                   30  
15          Lys Ile Leu Trp Ser Arg Ala Ile  
          35              40

20          <210> 154  
      <211> 25  
      <212> PRT  
      <213> Homo sapiens

25          <220>  
      <221> SITE  
      <222> (25)  
      <223> Xaa equals stop translation

30          <400> 154  
      Met Thr Pro Gly Leu Phe Leu Tyr Phe Val Cys Val Cys Val Ser His  
          1               5                   10               15  
          Cys Ala Gly Leu Gly Gln Leu Ser Xaa  
35          20              25

40          <210> 155  
      <211> 103  
      <212> PRT  
      <213> Homo sapiens

45          <400> 155  
      Ile Arg His Glu Leu Gly Cys Ser Trp Arg Phe Arg Ala Val Lys Ala  
          1               5                   10               15  
          Ala Ser Ala Gln Gly Leu Phe Leu Ser Ala Pro Gly Pro Ala Ala Arg  
          20              25                   30

50          Arg Cys His Gly Val Val Arg Cys Phe Ser Thr Cys Arg Ala Leu Thr  
          35              40                   45

55          Ala Arg Cys Thr Gly Arg Val Pro Trp Glu Ala Cys Leu Tyr Ser Ser  
          50              55                   60

60          Glu Pro Pro Leu Thr Glu Thr Val Ala Arg Ser Val Ser Trp Thr Cys  
          65              70                   75               80  
          Glu Leu Ala Leu Thr Cys Tyr Ala Pro Arg Ala Leu Ser Gly Ala Pro  
          85              90                   95

Val Leu Cys Arg His Asp Val  
100

5

<210> 156  
<211> 46  
<212> PRT  
<213> Homo sapiens

10

<400> 156  
Phe Leu Ala Ile His Phe Pro Thr Asp Phe Pro Leu Lys Pro Pro Lys  
1 5 10 15

15

Val Ala Phe Thr Arg Met Tyr Phe Pro Asn Ser Asn Ser Asn Gly Ser  
20 25 30

Thr Cys Leu Asp Ile Leu Trp Ser Gln Trp Ser Pro Ala Leu  
35 40 45

20

<210> 157  
<211> 101  
<212> PRT

25 <213> Homo sapiens

<400> 157  
Met Leu Leu Thr Pro His Phe Asn Val Ala Asn Pro Gln Asn Leu Leu  
1 5 10 15

30

Ala Gly Leu Trp Leu Glu Asn Glu His Ser Phe Thr Leu Met Ala Pro  
20 25 30

35

Glu Arg Ala Arg Thr His His Cys Gln Pro Glu Glu Arg Lys Val Leu  
35 40 45

Phe Cys Leu Phe Pro Ile Val Pro Asn Ser Gln Ala Gln Val Gln Pro  
50 55 60

40

Pro Gln Met Pro Pro Phe Cys Cys Ala Ala Ala Lys Glu Lys Thr Gln  
65 70 75 80

Glu Glu Gln Leu Gln Glu Pro Leu Gly Ser Gln Cys Pro Asp Thr Cys  
85 90 95

45

Pro Asn Ser Leu Cys  
100

50

<210> 158  
<211> 211  
<212> PRT  
<213> Homo sapiens

55

<400> 158  
Met Arg Leu Phe Leu Trp Asn Ala Val Leu Thr Leu Phe Val Thr Ser  
1 5 10 15

60

Leu Ile Gly Ala Leu Ile Pro Glu Pro Glu Val Lys Ile Glu Val Leu  
20 25 30

Gln Lys Pro Phe Ile Cys His Arg Lys Thr Lys Gly Gly Asp Leu Met  
           35                        40                        45

5   Leu Val His Tyr Glu Gly Tyr Leu Glu Lys Asp Gly Ser Leu Phe His  
       50                        55                        60

Ser Thr His Lys His Asn Asn Gly Gln Pro Ile Trp Phe Thr Leu Gly  
   65                        70                        75                        80

10   Ile Leu Glu Ala Leu Lys Gly Trp Asp Gln Gly Leu Lys Gly Met Cys  
       85                        90                        95

15   Val Gly Glu Lys Arg Lys Leu Ile Ile Pro Pro Ala Leu Gly Tyr Gly  
       100                      105                      110

Lys Glu Gly Lys Gly Lys Ile Pro Pro Glu Ser Thr Leu Ile Phe Asn  
   115                      120                        125

20   Ile Asp Leu Leu Glu Ile Arg Asn Gly Pro Arg Ser His Glu Ser Phe  
       130                      135                      140

Gln Glu Met Asp Leu Asn Asp Asp Trp Lys Leu Ser Lys Asp Glu Val  
   145                      150                        155                        160

25   Lys Ala Tyr Leu Lys Lys Glu Phe Glu Lys His Gly Ala Val Val Asn  
       165                      170                      175

30   Glu Ser His His Asp Ala Leu Val Glu Asp Ile Phe Asp Lys Glu Asp  
       180                      185                      190

Glu Asp Lys Asp Gly Phe Ile Ser Ala Arg Glu Phe Thr Tyr Lys His  
   195                      200                      205

35   Asp Glu Leu  
       210

<210> 159  
 40   <211> 186  
 <212> PRT  
 <213> Homo sapiens

<400> 159  
 45   Glu Val Lys Ile Glu Val Leu Gln Lys Pro Phe Ile Cys His Arg Lys  
       1                      5                      10                      15

Thr Lys Gly Gly Asp Leu Met Leu Val His Tyr Glu Gly Tyr Leu Glu  
   20                      25                      30

50   Lys Asp Gly Ser Leu Phe His Ser Thr His Lys His Asn Asn Gly Gln  
       35                      40                      45

55   Pro Ile Trp Phe Thr Leu Gly Ile Leu Glu Ala Leu Lys Gly Trp Asp  
       50                      55                      60

Gln Gly Leu Lys Gly Met Cys Val Gly Glu Lys Arg Lys Leu Ile Ile  
   65                      70                      75                      80

60   Pro Pro Ala Leu Gly Tyr Gly Lys Glu Gly Lys Gly Lys Ile Pro Pro



Ser Leu Leu Lys Ile Gln  
20

5      <210> 162  
      <211> 57  
      <212> PRT  
      <213> Homo sapiens

10     <220>  
      <221> SITE  
      <222> (34)  
      <223> Xaa equals any one of the naturally occurring L-amino acids

15     <400> 162  
Met Phe Tyr Asn Phe Val Arg Gln Leu Asp Thr Val Ser Ile Glu His  
1                5                10                15

20     Ala Gly Lys Ser Lys Leu Lys Met Thr Val Gly Thr Lys Leu Thr Ser  
20                25                30

25     Gly Xaa Gly Pro Arg Lys Ser Ser Gln Ser Gly Arg Ile Ala Ala Ser  
35                40                45

30     Ile Thr Asp Cys Gln Gln Cys Lys Ala  
50                55

35     <210> 163  
      <211> 46  
      <212> PRT  
      <213> Homo sapiens

35     <220>  
      <221> SITE  
      <222> (16)  
      <223> Xaa equals any one of the naturally occurring L-amino acids

40     <400> 163  
Met Glu Ala Ala Ile Leu Pro Leu Trp Leu Leu Phe Leu Gly Pro Xaa  
1                5                10                15

45     Pro Glu Val Ser Phe Val Pro Thr Val Ile Phe Asn Leu Asp Phe Pro  
20                25                30

45     Ala Cys Ser Ile Leu Thr Val Ser Ser Cys Leu Thr Lys Leu  
35                40                45

50     <210> 164  
      <211> 25  
      <212> PRT  
      <213> Homo sapiens

55     <400> 164  
Asn His Gly His Ser Cys Phe Leu Cys Glu Ile Val Ile Arg Ser Gln  
1                5                10                15

60     Phe His Thr Thr Tyr Glu Pro Glu Ala  
20                25

5 <210> 165

<211> 48

<212> PRT

<213> Homo sapiens

10 <400> 165

Ser Gly Arg His Arg Val Glu Leu Gln Leu Leu Phe Pro Leu Val Arg  
1 5 10 15

Val Asn Phe Glu Leu Gly Val Asn His Gly His Ser Cys Phe Leu Cys  
20 25 30

15 Glu Ile Val Ile Arg Ser Gln Phe His Thr Thr Tyr Glu Pro Glu Ala  
35 40 45

20

25 <210> 166

<211> 141

<212> PRT

<213> Homo sapiens

30 <400> 166

Met Asn Ala Arg Gly Leu Gly Ser Glu Leu Lys Asp Ser Ile Pro Val  
1 5 10 15

35 Thr Glu Leu Ser Ala Ser Gly Pro Phe Glu Ser His Asp Leu Leu Arg  
20 25 30

Lys Gly Phe Ser Cys Val Lys Asn Glu Leu Leu Pro Ser His Pro Leu  
35 40 45

Glu Leu Ser Glu Lys Asn Phe Gln Leu Asn Gln Asp Lys Met Asn Phe  
50 55 60

40 Ser Thr Leu Arg Asn Ile Gln Gly Leu Phe Ala Pro Leu Lys Leu Gln  
65 70 75 80

Met Glu Phe Lys Ala Val Gln Gln Val Gln Arg Leu Pro Phe Leu Ser  
85 90 95

45 Ser Ser Asn Leu Ser Leu Asp Val Leu Arg Gly Asn Asp Glu Thr Ile  
100 105 110

50 Gly Phe Glu Asp Ile Leu Asn Asp Pro Ser Gln Ser Glu Val Met Gly  
115 120 125

Glu Pro His Leu Met Val Glu Tyr Lys Leu Gly Leu Leu  
130 135 140

55

<210> 167

<211> 15

<212> PRT

60 <213> Homo sapiens

<400> 167  
Gly Gly Asn Lys Tyr Gln Thr Ile Asp Asn Tyr Gln Pro Tyr Pro  
1 5 10 15

5 <210> 168  
<211> 20  
<212> PRT  
<213> Homo sapiens

10 <400> 168  
Pro Leu Leu Gly Val Ser Ala Thr Leu Asn Ser Val Leu Asn Ser Asn  
1 5 10 15

15 Ala Ile Lys Asn  
20 <210> 169  
<211> 14  
<212> PRT  
<213> Homo sapiens

25 <400> 169  
Gly Ser Ala Val Ser Ala Ala Pro Gly Ile Leu Tyr Pro Gly  
1 5 10

30 <210> 170  
<211> 81  
<212> PRT  
<213> Homo sapiens

35 <400> 170  
Ala Gly Ile Gln His Glu Leu Ala Cys Asp Asn Pro Gly Leu Pro Glu  
1 5 10 15

Asn Gly Tyr Gln Ile Leu Tyr Lys Arg Leu Tyr Leu Pro Gly Glu Ser  
20 25 30

40 Leu Thr Phe Met Cys Tyr Glu Gly Phe Glu Leu Met Gly Glu Val Thr  
35 40 45

Ile Arg Cys Ile Leu Gly Gln Pro Ser His Trp Asn Gly Pro Leu Pro  
45 50 55 60

Val Cys Lys Val Ala Glu Ala Ala Glu Thr Ser Leu Glu Gly Gly  
65 70 75 80

50 Asn

55 <210> 171  
<211> 27  
<212> PRT  
<213> Homo sapiens

60 <400> 171  
Gln Pro Ser His Trp Asn Gly Pro Leu Pro Val Cys Lys Val Ala Glu

1                   5                   10                   15

Ala Ala Ala Glu Thr Ser Leu Glu Gly Gly Asn  
20                   25

5

&lt;210&gt; 172

&lt;211&gt; 13

&lt;212&gt; PRT

10   &lt;213&gt; Homo sapiens

&lt;400&gt; 172

Tyr Glu Thr Gly Glu Thr Arg Glu Tyr Glu Val Ser Ile  
1                   5                   10

15

&lt;210&gt; 173

&lt;211&gt; 195

&lt;212&gt; PRT

20   &lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (40)

25   &lt;223&gt; Xaa equals any one of the naturally occurring L-amino acids

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (161)

30   &lt;223&gt; Xaa equals any one of the naturally occurring L-amino acids

&lt;400&gt; 173

Asp Asp Asp Gly Leu Pro Phe Pro Thr Asp Val Ile Gln His Arg Leu  
1                   5                   10                   15

35

Arg Gln Ile Glu Ala Gly Tyr Lys Gln Glu Val Glu Gln Leu Arg Arg  
20                   25                   30

40

Gln Val Arg Asp Ser Asp Glu Xaa Gly His Pro Ser Leu Leu Cys Pro  
35                   40                   45Ser Ser Arg Ala Pro Met Asp Tyr Glu Asp Asp Phe Thr Cys Leu Lys  
50                   55                   60

45

Glu Ser Asp Gly Ser Asp Thr Glu Asp Phe Gly Ser Asp His Ser Glu  
65                   70                   75                   80Asp Cys Leu Ser Glu Ala Ser Trp Glu Pro Val Asp Lys Lys Glu Thr  
85                   90                   95

50

Glu Val Thr Arg Trp Val Pro Asp His Met Ala Ser His Cys Tyr Asn  
100                   105                   110

55

Cys Asp Cys Glu Phe Trp Leu Ala Lys Arg Arg His His Cys Arg Asn  
115                   120                   125Cys Gly Asn Val Phe Cys Ala Gly Cys Cys His Leu Lys Leu Pro Ile  
130                   135                   140

60   Pro Asp Gln Gln Leu Tyr Asp Pro Val Leu Val Cys Asn Ser Cys Tyr

	145	150	155	160
	Xaa Thr His Ser Ser Leu Ser Cys Gln Gly Thr His Glu Pro Thr Ala			
	165	170	175	
5	Glu Glu Thr His Cys Tyr Ser Phe Gln Leu Asn Ala Gly Glu Lys Pro			
	180	185	190	
10	Val Gln Phe			
	195			
	<210> 174			
	<211> 28			
15	<212> PRT			
	<213> Homo sapiens			
	<400> 174			
20	Ser Glu Ala Ser Trp Glu Pro Val Asp Lys Lys Glu Thr Glu Val Thr			
	1	5	10	15
	Arg Trp Val Pro Asp His Met Ala Ser His Cys Tyr			
	20	25		
25				
	<210> 175			
	<211> 10			
	<212> PRT			
	<213> Homo sapiens			
30				
	<400> 175			
	His His Cys Arg Asn Cys Gly Asn Val Phe			
	1	5	10	
35				
	<210> 176			
	<211> 14			
	<212> PRT			
	<213> Homo sapiens			
40				
	<400> 176			
	Arg Leu Arg Gln Ile Glu Ala Gly Tyr Lys Gln Glu Val Glu			
	1	5	10	
45				
	<210> 177			
	<211> 87			
	<212> PRT			
	<213> Homo sapiens			
50				
	<400> 177			
	Met Ser His Cys Ala Arg Pro Leu Phe Phe Glu Thr Phe Phe Ile Leu			
	1	5	10	15
55	Leu Ser Pro Arg Leu Lys Cys Ser Gly Thr Asn Thr Val His Tyr Ser			
	20	25	30	
	Leu Asp Leu Leu Gly Ser Ser Asn Ser Ala Ser Val Pro Gln Val Gly			
	35	40	45	
60				

80

Gly Leu Thr Asn Ala Gln His Asp Thr Trp Leu Ile Phe Val Phe Cys  
50 55 60

5 Val Cys Val Cys Glu Pro Leu Arg Arg Pro Trp Ala Ala Phe Leu Ile  
65 70 75 80

Ser Val Thr Ser Ser Ile Lys  
85

10

&lt;210&gt; 178

&lt;211&gt; 30

&lt;212&gt; PRT

15

&lt;213&gt; Homo sapiens

20

<400> 178  
Val Pro Gln Val Gly Gly Leu Thr Asn Ala Gln His Asp Thr Trp Leu  
1 5 10 15

Ile Phe Val Phe Cys Val Cys Val Cys Glu Pro Leu Arg Arg  
20 25 30